PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 15 February 15, 1929 Number 2

CONTROL OF SEX IN CLADOCERA. III. LOCALIZATION OF THE CRITICAL PERIOD FOR CONTROL OF SEX[1]

By Arthur M. Banta and L. A. Brown

Carnegie Institution of Washington, Cold Spring Harbor, N. Y., and Zoölogical Laboratory, State University of Iowa and Carnegie Institution of Washington

Communicated January 12, 1929

Every young female of *Moina macrocopa* has the possibility of producing in the first adult instar either of two types of eggs. She may produce (1) sexual eggs, or (2) parthenogenetic eggs. The former are haploid and do not develop unless fertilized. The latter are diploid and develop at once into females or males. The normal or usual course is the production of parthenogenetic females. Sexual eggs and males are called forth by very special and quite diverse environmental conditions.[2,3,4] The experimenter can, by appropriate treatment during the period of development, cause an animal, which would normally produce female progeny, to produce males. It is obvious, then, that the sex of the forthcoming parthenogenetic eggs is undetermined during the early growth stages of the mother, but definitely fixed at some period before the eggs are laid.

Before crucial work on the control of sex in this species could be done, it was necessary to determine with some exactness the position, in terms of the developmental stages of the mother, of this critical period in the development of the eggs during which they are amenable to control of sex. The present paper presents some of these data (for first broods only) for one species of Cladocera, *Moina macrocopa*.

Experimental.—Since the length of the developmental period of *M. macrocopa* varies greatly with temperature, having a higher temperature coefficient (Q_{10} = 2.39, for the temperature interval 20° to 30°C.) than most common Cladocera,[5] it was necessary to have definite morphological units as an index of the age, rather than time units. Since the number of preadult instars of a daphnid is constant, regardless of the temperature, and since the different instars present marked differences in length of the animals, even though lacking strictly structural dissimilarity, the instar was chosen as the index of the age of the animals. The young females of *M. macrocopa* when released from their mother's brood chamber have an

average length of 0.591 mm. Following the first moult, the length of the female increases to 0.717 mm. The third instar animal (following the second ecdysis) has an average length of 0.892 mm. During the fourth instar, at the beginning of which the first clutch of eggs passes from the ovaries to the brood chamber, the female has an average length of 1.090 mm. There are, then, two juvenile instars, and one adolescent instar before the first adult instar; and these can be distinguished readily by size. The average lengths for the first, second, third and fourth (first adult) instars were based on 333, 256, 387 and 182 measurements, respectively. The durations of the instars are not equal. At 20°C. the length of life to maturity (first three instars) is about 65 hours; 38 hours at 25°; and 30 hours at 30°. The fourth instar, at the beginning of which the parthenogenetic eggs pass into the brood chamber and at the end of which the young are released, is equal to approximately 70 per cent of the combined length of the three earlier instars. Hence a young female at 20°C. produces her first young when about 114 hours old; at 30°C. this time is in the neighborhood of 50 hours.

While the length of the animals does not change appreciably during an instar, certain differences in the appearance of the ovary are apparent. These changes in the ovary are slight during the first and second instars, and consist mainly in the relative, though slight, increase in size of the ovary and the increasing prominence of the egg nuclei. But during the third instar the changes are more conspicuous. The nuclei are plainly visible during the first two-thirds of this instar but during the middle third they are becoming obscured by the deposition of yolk granules within the eggs. During the latter part of this instar the nuclei are completely hidden by the large amount of yolk present. These changes in the amount of yolk in the eggs are, of course, accompanied by changes in the size and shape of the entire ovary, which becomes distended and lobed due to the increased mass of the contained eggs. Unfortunately in this species the yolk is not highly colored as in some other species, and so does not give well-defined landmarks within the instar.

It was known from some preliminary experiments,[2] using elapsed time as an index of the age of the animals, that the sex of the developing eggs was not definitely fixed until quite late in the development of the mother. These experiments were conducted as follows: One or more individuals of a large brood (20 or more) of newly released parthenogenetic young were isolated in bottles containing 75 cc. of culture water. The remaining young of the same brood were placed in a single bottle containing the same amount of culture water. It was known from earlier work[2,3] that females crowded together would produce a high percentage of males. In the present experiments animals were removed at various times from the crowded bottle and isolated in the usual amount of culture medium.

Mothers not crowded at all, and mothers crowded for any length of time up to 31 to 46 hours (17° to 19°C.) and then isolated, failed to produce males. Mothers crowded 48 to 76 hours gave 45.0 per cent males.

TABLE 1

Experiments to localize the critical period previous to which control measures are ineffective.

Mothers crowded and later isolated at definite times to determine at how late a period the crowding must exist in order to be effective. The numbers of mothers isolated at each hour period are in parentheses. The "numbers of hours" were determined by frequent examination of the mothers used in these experiments to determine just when the eggs were laid. Production of males by mothers (in each case by a single mother) isolated 10, 8 and 6 hours, respectively, before the eggs were laid is exceptional.[7] Since these experiments were conducted at 25–26°C. and since an adult instar at this temperature occupies about one half the time required at 20°C. the times in hours given in this table have been multiplied by two so as to make the times directly comparable with those obtained at 20° as given in Table 2.

NUMBER OF HOURS BEFORE EGG LAYING WHEN MOTHERS WERE ISOLATED, I.E., WHEN CROWDING CEASED	EXPERIMENT NO. 2826 SEX OF YG. ♀ ♂			2827 SEX OF YG. ♀ ♂			2828 SEX OF YG. ♀ ♂			2832 SEX OF YG. ♀ ♂			TOTAL YOUNG SEXED ♀ ♂		AVE. % ♂
	(5)						(2)			(2)			(9)		
12	11	0					5	0		17	0		33	0	0
				(1)			(1)			(4)			(6)		
11				5	0		3	0		39	0		47	0	0
							(4)			(2)			(6)		
10							14	5		14	0		28	5	15
							(2)			(2)			(4)		
9							12	0		13	0		25	0	0
															5
	(1)			(1)			(3)			(2)			(7)		
8	4	0		7	0		12	5		14	0		37	5	12
				(2)			(6)			(1)			(9)		
7				13	0		26	0		6	0		45	0	0
							(4)			(3)			(7)		
6							19	0		12	6		31	6	16
	(1)			(2)			(4)			(2)			(9)		
5	3	0		16	0		19	0		10	0		48	0	0
	(3)			(2)			(6)						(11)		
4	1	10		4	11		20	4					25	25	50
	(4)			(5)			(2)			(2)			(13)		
3	7	6	69 % ♂'s	18	12	35 % ♂'s	0	7	66 % ♂'s	8	0	20 % ♂'s	33	25	43
				(4)			(6)			(9)			(19)		47.5
2				20	8		12	17		31	11		63	36	36
	(4)			(2)			(11)			(3)			(20)		
1	3	9		16	0		8	51		17	3		44	63	59
0 (not isolated)	29	78	73 % ♂'s	191	69	27 % ♂'s	177	259	59 % ♂'s	102	51	33 % ♂'s	499	457	47.8
													958	622	
													1580		

Mothers not isolated but left in the crowded bottles until after the eggs of the first clutch were laid (72 to 96 hours) gave 34.1 per cent males among the 1066 young sexed. These experiments[6] while not significant as to time because of the lack of temperature control, indicated that the sex of the eggs is not determined until late in their ovarian development.

A second series of experiments, involving 5995 sexed young, was conducted in a similar manner, but the age of the animals was determined by attained instars, which made the record independent of the temperature. Females from the same brood, having been originally heavily crowded and hence if undisturbed destined to produce a large percentage of males, were isolated, several at each of the following stages—first instar, second instar, early third instar, late third instar (a short time before the eggs were laid), and early fourth instar (a short time after the eggs were laid). A number remained as controls, i.e., were not isolated at all. These age classes gave, respectively, the following percentages of males: 0, 0, 0, 9.0, 57.6 and 41.2. These data localized the critical period for the control of sex near the end of the adolescent instar. The 9.0 per cent males produced by those isolated in the latter part of the third instar were probably produced by mothers which were actually near the end of the third instar.

Attention was now directed toward a more critical analysis of the time relations near the end of the third instar. While, as indicated above, there are some morphological changes in the developing eggs due to the deposition of yolk, these are not marked enough to serve as landmarks in critically determining the stage of development of the animals within this instar so that time intervals had to be used again. But as the time periods were short, the temperature variations were slight. These later series of experiments were conducted in like manner to those just discussed, but the isolations were made at definite time periods during only the latter half of the third instar; and the actual time of egg-laying was determined by occasional microscopic examination of some of the mothers. A summary of the results of one of these series is shown in Table 1.

Except for three exceptional occurrences of males,[7] no males were produced by the 57 mothers isolated more than four hours before egg-laying. Those isolated 4 hours before egg-laying produced 50 per cent males; those isolated 3, 2 and 1 hours before egg-laying produced 43, 36 and 59 per cent males, respectively. The average male production for all those mothers isolated 4 or less than 4 hours before egg-laying was 47.5 per cent; while those mothers not isolated at all produced substantially the same male percentage, 47.8 per cent. Other experiments showed similar results. It is clear that isolation of mothers more than 4 hours before egg-laying (at 20° to 22°C.) ordinarily prevents male production by previously crowded mothers; but that after this time isolation is without effect upon the sex of the forthcoming young.

Other experiments were planned in such a way that the animals which had previously been under conditions in which only female production ordinarily occurs could be subjected almost immediately to an environment which influences male production, i.e., crowding, instead of relieving them of this condition, as in the series of experiments just reported. This was done in the following manner. A double row of bottles was prepared in advance. Into one row of bottles was placed a large brood of animals, three (occasionally only two) animals per bottle; each of the bottles in the second row contained a large population of animals of the same species. At each successive hour during the last nine to twelve hours of the third

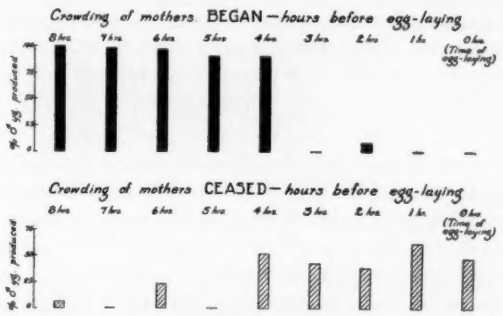


FIGURE 1

A diagrammatic presentation of the results of the series of experiments tabulated in Tables I and II. The shaded series of polygons show the results of a series of experiments in which the crowding was less effective than in the series represented by the black polygons, hence the shaded polygons are less tall than the black polygons. The one small black polygon and the two small shaded polygons indicate the appearance of small numbers of males where they ordinarily do not occur.

instar the three animals previously isolated in one of the first row of bottles were placed in the corresponding crowded bottles. These test mothers were confined within a section of glass tubing, the lower end of which was covered by a piece of silk bolting cloth to keep the animals in and to allow free interchange of the excretory products between the animals outside and those inside the tube. After all of the animals had reached the fourth instar and the eggs were laid, they were returned to their original bottles and kept until the first brood of young was released. These young were then sexed. Table 2 gives the detailed data for three such experiments. The average male production of the 68 mothers crowded until after egg-laying, from times beginning from 12 to 4 hours before egg-laying, is 77.9

per cent. The percentages of males for the different time classes for all three experiments taken together show a very striking change between four and three hours previous to egg laying. The eight animals subjected to crowding for four hours gave 82 per cent males, while the six animals crowded for only three hours before egg laying failed to produce any males

TABLE 2

Experiments to localize the critical period beyond which sex is no longer subject to control measures.

Mothers reared isolated and then subjected to crowded conditions, beginning at definite times, to determine how early the crowding must exist in order to be effective. The "numbers of hours" were determined by examination of control females (from the brood of mothers used in each experiment) to learn just when their eggs were laid. Some variation in actual time of egg-laying occurs. Presumably the female in Experiment 543 which was crowded for only two hours before the eggs were supposed to have been laid, and yet produced some males, was actually four hours or more from her actual time of laying when first subjected to the crowded conditions.

NUMBER OF HOURS BEFORE EGG LAYING WHEN CROWDING BEGAN	EXPERIMENT NO. 541 SEX OF YG. ♀	♂		542 SEX OF YG. ♀	♂		543 SEX OF YG. ♀	♂		TOTAL YOUNG SEXED ♀	♂	AVE. % ♂'s	
12	(3) 0	19					(2) 17	0		(5) 17	19	53	
11	(3) 19	16					(2) 2	14		(5) 21	30	59	
10	(3) 0	23		(3) 22	0		(3) 2	21		(9) 24	44	65	
9	(3) 0	18		(3) 23	9		(3) 25	1		(9) 48	28	35	
8	(3) 0	19	92 % ♂'s	(3) 0	33		(3) 0	27	73 % ♂'s	(9) 0	79	100	77.9
7	(3) 0	29				69 % ♂'s	(3) 1	25		(6) 1	54	98	
6	(3) 0	32		(3) 5	20		(3) 0	58		(9) 5	110	96	
5	(2) 0	22		(3) 0	27		(3) 17	30		(8) 17	79	82	
4	(2) 0	31		(3) 10	42		(3) 13	33		(8) 23	106	82	
3	(2) 29	0	0 % ♂'s	(2) 35	0	0 % ♂'s	(2) 32	0	18 % ♂'s	(6) 96	0	0	4.8
2	(2) 29	0		(6) 93	0		(2) 20	12		(10) 142	12	8	
										394	561		
										855			

at all.[8] This result, together with the results of experiments discussed in earlier paragraphs of this paper and of other series of experiments not here recorded, clearly indicates that the sex of the developing egg is subject to control only during a period localized at about four hours (at 20°C.) before the time it leaves the ovary and passes to the brood chamber. Figure 1 shows diagrammatically the results of the series of experiments tabulated in tables 2 and 3.

Discussion.—The parthenogenetic egg of cladocerans is diploid whether the egg produces a female individual[9,10] or a male.[10,11,12] In *Moina* the first and only maturation division in the parthenogenetic egg occurs just after the egg is laid and there is no reduction in the number of chromosomes.[11,13] Allen's findings indicate that the spindle and the small chromosomes appear shortly (probably an hour or less) before the eggs are laid. We have a situation here, then, in which the sex of the parthenogenetic egg is fixed by external environmental conditions effective for a period ending (at 20°C.) about three and a half hours before the chromosomes appear or the spindle of the single maturation division is formed.

It has so far proved impossible to determine as definitely as we should like the time of fixation (for sectioning and cytological study) of a given mother with reference to the time at which she would have laid her eggs. The actual time of egg-laying of sisters reared in the same culture bottle was learned and used as a means of calculating the stage at which fixations were presumably made. But these sisters do not all lay their eggs at once. In fact, their times of egg-laying may vary over a range of as much as four hours. In preparing material upon which Dr. Allen's[10] cytological observations were made the average time of egg-laying of the control sisters was assumed to be the time at which those fixed would have laid their eggs. But obviously such staging is only approximate. Further, this is very difficult cytological material. Few of the egg-bearing mothers which are fixed yield cytologically useful material. These few were of the calculated stage of 1 hour or less before egg-laying. The fact that spindle formation was found only in cases in which the time of fixation was presumably about $^1/_2$ hour before egg-laying, and not in the presumably earlier stages which Dr. Allen studied, inclines the authors to believe that the beginning of spindle formation is probably actually near the calculated time—about $^1/_2$ hour before egg-laying. It is still possible, however, that this event begins to occur earlier, i.e., nearer the experimentally determined critical period for sex-control (approximately 4 hours before the eggs are laid) than the present observations indicate. To the authors it seems improbable that the localization of the critical period can be narrowed down much further or that, on the other hand, the stage at which animals are killed for cytological examination can be much more precisely determined.

Another consideration is of interest with reference to the time relation between the critical period for sex-control and the beginning of the maturation division. The experimentally determined critical period (about 4 hours before the eggs are to be laid) is, we think, approximately correct for the *external* environment. But we have no information nor do we know any feasible means of obtaining precise information concerning the the location of the critical period for the *internal* environment of the ovarian eggs. With the apparent time of spindle formation located at about $^1/_2$ hour before the eggs are to be laid and about $3^1/_2$ hours after the critical period for the external environment, it is still possible that the critical period for the internal environment coincides with or immediately precedes the beginning of spindle formation. For it is possible that the internal environment retains the effects of crowding for $3^1/_2$ hours after the actual crowding has ceased; and conversely that only in animals which have been crowded as much as (or more than) $3^1/_2$ hours does the internal environment reach sufficient saturation of the effects of crowding to affect the maturation division. If this supposition is correct, with reference to the *internal* environment, the critical period may readily be about $^1/_2$ hour before the eggs are to be laid and thus coincide with the time of initiation of the maturation division.

The fact that external environmental influences operative *previous* to the formation of the maturation spindle control the sex of *Moina* eggs is perhaps consistent with any of the following interpretations: (1) A direct action of the environment on the egg and not through a chromosome mechanism. (2) An effect through the directing of a sex chromosome complex. (3) An effect through a more deeply seated and fundamental mechanism to which the sex chromosome mechanism is subordinate.

1. A direct action of the environment on the egg, and not on the behavior of sex chromosomes during the maturation division, might seem to apply. But it is difficult to understand how this action can be accomplished without the more frequent appearance of sex intergrades (only one having been recognized in our *Moina* material). The extent of the crowding, i.e., the concentration of the influential excretory products, acts to give varying numbers of males, but the males so produced are normal; and the sex intergrades that have appeared in cladocerans, so far as critically studied, behave as mendelian dominants[14,15] in sexual reproduction and so are clearly genic in inheritance.

A further difficulty with this type of explanation is the wide, almost universally demonstrated occurrence of the sex-chromosome mechanism in the different groups of animals. It seems improbable that Cladocera, at least eight species of which are subject to sex control by environmental means,[2] lack such a mechanism; and it seems improbable that such a mechanism, if present, would be generally over-ridden by environmental

influences in even so specialized a group as the Cladocera, unless indeed there is a mechanism more fundamental than the sex chromosome mechanism.

2. Inasmuch as the cytology of this group has not been exhaustively studied, one is left free to suggest that the manner of sex-determination in *Moina* may not be greatly different from that found by Morgan[16] in phylloxerans and aphids in which the parthenogenetic males are diploid, their chromosome complement being the result of the elimination during maturation in oögenesis of half the X-chromosomes. The male so produced gives only female-determining sperm inasmuch as the sperm containing no X-chromosomes degenerate. Certainly Cladocera, like the phylloxerans and aphids, have diploid parthenogenesis and the males so produced are diploid. The fertilized egg produces only female young. There is a suggestion in the paper of Chambers[17] on *Simocephalus* that only a portion of the sperm (presumably the female determiners) survive, but of this Chambers was not convinced. But so far no studies are at hand to indicate that the male has a chromosome complement different from that of the female. The only present indication from the cytological side is Allen's[10] report of 22 as the chromosome count for eggs produced by *Moina macrocopa* mothers under conditions which ordinarily cause them to produce a high proportion of males. If this chromosome number applies to the male, as well as the female, then the interpretation suggested above will not apply unless the sex-determining mechanism is of the WZ type. But this latter suggestion involves such additional assumptions (without known cytological precedent), in order to harmonize with the facts of the life cycle, that it fails to retain plausibility.[18]

Schrader[19] has recently reviewed the cytological information concerning the sex chromosomes in crustaceans. He cites evidence that in two Malacostraca, *Cambarus* and *Gammarus*, there are sex chromosomes, with indications that the male is the heterozygous sex. However, with the copepods, constituting an order of animals more nearly related to Cladocera, the somewhat fragmentary cytological evidence suggests that in some species the female is the heterozygous sex and has the XO-condition. Should the female Cladoceran prove to be heterozygous for sex, with the XO-condition, the cytological implications of our results might be somewhat as follows. The usual parthenogenetic Cladoceran egg would then presumably have, as its one maturation division, an equational division, the egg retaining the full autosomal set of chromosomes and the one X-chromosome and would produce a female. The parthenogenetic egg destined to become a male would, on the other hand, have to achieve a duplication of the X-chromosome. In this case the sperm should be of one sort and, in order to harmonize with the fact that the fertilized sexual egg always produces a female, one must assume that sexual eggs of

only one sort, without an X-chromosome, are produced (or are hatchable).

If a sex chromosome complex is present in Cladocera, either of the above theoretical explanations would appear to fit the observed facts but, except for the first mentioned hypothesis, they involve somewhat heterodox assumptions and, so far as Cladocera cytology is concerned, none of them have any supporting evidence. Any type of explanation which assumes a chromosome mechanism as a basis for sex-determination in Cladocera must take into account the fact that something in the mother's environment influences the distribution of the sex-chromosomes in the one (non-reductional) maturation division of the parthenogenetic egg.

3. Numerous workers have suggested or implied an inadequacy of the sex chromosome hypothesis. It is theoretically possible that there may be, in our material and elsewhere, a mechanism more fundamental than the sex chromosome complex to which the latter (very wide and possibly universal in its occurrence) is subservient and of which it is merely the visible index.

Whatever the theoretical explanation, the parthenogenetic egg of *Moina macrocopa* may develop into either a female or a male, depending upon environmental conditions and the environmental factors which control the determination of sex in *Moina macrocopa* are effective during a period which antedates the formation of the maturation spindle and the appearance of definitive chromosomes apparently by more than three hours (20° to 23°C.).

The knowledge of the location of the critical period was of extreme value in later experiments in which agencies other than crowding were used in controlling sex.

Summary.—1. *Moina macrocopa* has two juvenile instars, and one adolescent instar, before it becomes adult. The first clutch of eggs passes from the ovary to the brood chamber within the first few minutes after the beginning of the first adult instar.

2. The critical period for the control of sex is about four hours (20°C.) before the eggs are laid.

3. Possible mechanisms of sex-determination in this species are discussed.

[1] From the Department of Genetics, Carnegie Institution of Washington, Cold Spring Harbor, N. Y.

[2] Banta, A. M., and Brown, L. A., "Eugenics, Genetics and the Family," *Sci. Papers 2nd Int. Cong. Eugenics*, **1**, 142 (1923).

[3] Banta, A. M., and Brown, L. A., *Physiol. Zoöl.*, **2**, 80 (1929).

[4] Banta, A. M., *Zeit. ind. Abs. Vererb.*, **40**, 28 (1925).

[5] Brown, L. A., *Proc. Soc. Exp. Biol. Med.*, **25**, 164 (1927).

[6] Tabulated in Banta and Brown, 1923, Table 3.

[7] Three unexpected cases of production of males (in each case by a single mother) occurred in Experiments 2828 and 2832 (Table 1). In one of these experiments the mother was isolated 10, in the other, 8 and 6 hours before egg laying. All the mothers isolated in these experiments were placed in vials to permit ready observation of the time of release of the young. In view of the very great rarity of male production by mothers isolated at such times it is believed that the mothers in these experiments were isolated in too small a quantity of culture medium and that consequently they "crowded" themselves, i.e., were influenced by their own excretory products sufficiently to induce male production in these three cases. Earlier experiments (Banta and Brown, 1929) had shown that individual mothers reared in 7.5 cc. of culture medium gave normal male production for crowded mothers. Other experiments conducted precisely like those tabulated in Table 1 confirmed our opinion that the unexpected male production by these three mothers was not due to the effects of the earlier crowding.

If one prefers not to accept the male production by three of the 57 mothers isolated more than four hours before egg-laying as exceptional, it is still clear that there is a vast difference in male production between mothers isolated more than four hours before egg-laying (produced 5 per cent males) and those isolated four or less than four hours before egg-laying (produced 47.5 per cent males). The location and significance of the critical period remains unaltered.

[8] The one mother, which produced some males (12) after supposedly having been crowded for only two hours before egg-laying, presumably was an irregular one which was actually four or more hours from egg-laying when first subjected to the crowding effect.

[9] Weismann, A., *Zoöl. Anz.*, **9**, 570 (1886).

[10] Allen, E., *Science, N. S.*, **67**, 18 (1928).

[11] Taylor, M., *Zoöl. Anz.*, **45**, 21 (1914).

[12] Banta, A. M., and Wood, T. R., *Science, N. S.*, **67**, 18 (1928).

[13] Kühn, A., *Arch. Zellforsch.*, **1**, 538 (1908).

[14] Banta, A. M., Snider, K., and Wood, T. R., *Proc. Soc. Exp. Biol. Med.*, **23**, 621 (1926).

[15] Banta, A. M., and Wood, T. R., *Verh. V. Int. Kong. Vererb., Zeit. ind. Abs. Vererb., Sup.-Bd.*, **1**, 391 (1928b).

[16] Morgan, T. H., *Jour. Exp. Zoöl.*, **19**, 285 (1915).

[17] Chambers, E., *Biol. Bull.*, **25**, 134 (1913).

[18] If the WZ type is assumed as the condition, the female is WZ and the male ZZ. Parthenogenetically produced females are WZ. But to get the parthenogenetically produced males a differential maturation involving the passing of the W to the polar body and a reduplication of the Z, or a transformation of the W to the Z type, would be necessary in order to arrive at the ZZ condition. Further, since all sexual eggs hatch as females, the mature sexual eggs should contain only W to account for which another differentiatial maturation must be assumed.

[19] Schrader, F., *The Sex Chromosomes*, Berlin, Borntraeger (1928).

GENETIC IDENTIFICATION OF THE SEX CHROMOSOMES IN SCIARA (DIPTERA)[1]

BY CHAS. W. METZ AND SILKA S. ULLIAN

DEPARTMENT OF GENETICS, CARNEGIE INSTITUTION OF WASHINGTON, COLD SPRING HARBOR, N. Y.

Communicated January 10, 1929

The peculiar relationship between the chromosomes of the two sexes in Sciara[2] and the peculiar chromosome behavior during spermatogenesis,[3] together with the regular production of "unisexual" progenies, due to the presence of male-producing and female-producing females,[4] have presented a series of phenomena difficult to harmonize on any simple scheme of sex determination. Perhaps the greatest difficulty in this respect is presented by the two large "sex limited" chromosomes, found only in males,[5] but apparently transmitted by all sperms. These chromosomes differ from the others in size and shape and remain condensed during the growth period of the primary spermatocyte in the manner typical of sex chromosomes. Thus their form and cytological behavior would indicate that they are sex chromosomes, but their distribution is such as to make it difficult to outline any simple scheme of sex determination based on the assumption that they are the primary sex determiners.

Without attempting a detailed discussion of the question of sex determination at this time, we present the accompanying summary of data which, we believe, demonstrate the presence of a pair of sex chromosomes distinct from the "sex limited" chromosomes just mentioned. These sex chromosomes are present in both males and females—the males being XY and the females XX in constitution.

The evidence is based on the inheritance of "swollen" wing veins, a clear-cut recessive mutant character. This character has been followed through more than ten generations and has shown typical sex-linked inheritance, except in one experiment discussed below and here the disturbance appears to be superimposed on the regular type of inheritance.

Swollen was first observed (by Ullian) in four males, the only offspring from a mass mating. These males were mated to wild-type virgin females from a related culture. From this mass mating 9 wild-type females and 7 wild-type males were obtained. The females were divided into three lots, which were outcrossed, two to males from wild-type stock and one to males from a related culture. The three matings gave the following results, in the order named:

(1) 84 wild-type females, 14 wild-type males and 10 swollen males;
(2) 144 wild-type females, 2 wild-type males and 1 swollen male;
(3) 55 wild-type females, 27 wild-type males and 34 swollen males.

In interpreting these results it must be remembered that individual females give "unisexual" progenies[5] and that the matings are mass matings. The ratio of females to males, therefore, is of no significance. Making allowance for this fact, the results are those expected on the basis of ordinary sex-linked inheritance—i.e., half the sons but none of the daughters are swollen.

Later matings involving heterozygous females by swollen males gave wild-type and swollen flies of both sexes as expected. But during this period and for some time afterward poor conditions in the cultures made the counts very small and unsatisfactory. Furthermore, the stocks used at that time were so interrelated that none was known definitely to be free from swollen. Consequently we may omit the records from these experiments (although they consistently indicate sex linkage) and consider only the subsequent ones in which such difficulties were eliminated.

In the latter experiments virgin swollen females were outcrossed singly to males from a wild stock just secured in California.[6] From such matings the sons were all swollen and the daughters all wild-type, as expected.

Since in this material "unisexual" progenies are obtained, the males and females come mainly from different parents. Seven mothers gave 212 wild-type daughters and seven others gave 355 swollen sons and 9 wild-type "exceptional" daughters. One bilateral gynandromorph appeared in which the male side was swollen and the female side wild-type.

Twenty-five F_1 females from matings of this kind were then tested individually by mating to swollen males from the same cross. Ten of these were "female-producers" and gave a total of 526 wild-type females, 1 swollen female, 7 wild-type males and 6 swollen males. The remaining fifteen were "male producers" and each gave both wild-type and swollen sons—total 261 wild-type and 369[7] swollen sons and one wild-type "exceptional" daughter.

At the same time five of the F_1 females were mated to wild-type males. One gave only wild-type daughters; the other four all gave only sons, both wild-type and swollen.

In the above tests of F_1 females the male progenies all give the results expected on the basis of ordinary sex-linked inheritance—i.e., the offspring are approximately half swollen and half wild-type. But in the female progenies, where the mothers were mated to swollen males and therefore should give both types of offspring, the expected swollen class is absent. Only one such female appeared where approximately 500 would be expected. This is the exceptional behavior mentioned above. That the mothers here were all actually heterozygous for swollen is made practically certain by two facts: first, that their male-producing sisters showed such a constitution and second, that six of the ten female-producers themselves gave "exceptional" swollen sons.

It is possible that the absence of swollen females in this case is due to poor viability, for such females have proved to be consistently less viable than the wild-type throughout our experiments. However, no such extreme inviability as this has been shown in any other experiment. Another possibility is that genetically swollen females were present here in the expected proportion, but failed to manifest the character somatically. This explanation is practically ruled out, however, by the fact that in other experiments no such behavior has been indicated and also by genetic tests of the females in question, as shown below.

Twenty-six of these females were tested individually by mating to swollen males. Fifteen were female-producers, all of which gave both wild-type and swollen daughters (total 433 wild-type, 218 swollen) thus showing that they were heterozygous. The remaining eleven, which were male-producers, gave similar results (total 389 wild-type, 349 swollen) except in one case. In the latter the eighteen offspring were all swollen. The mother may, therefore, have been homozygous for swollen in this case, although the presence of a sex-linked lethal might account for the result equally well.

In succeeding generations from this same series the inheritance has been that of a typical sex-linked character. For instance, in the generation following that just considered, twenty heterozygous female-producing females were tested and all gave both wild-type and swollen daughters. In four of these cases the actual numbers were not recorded. The other sixteen gave a total of 259 wild-type females, 120 swollen females, 4 wild-type males and 8 swollen males. In the next generation forty-four heterozygous female-producers were tested, with similar results, except in one case where only sixteen offspring were secured. The heterozygous male-producing females in these experiments have consistently given both wild-type and swollen sons, as expected, regardless of the type of males to which they were mated. Likewise, from outcrosses of swollen females to wild-type males the daughters have regularly been all wild-type and the sons all swollen, as expected, and from outcrosses of swollen males to wild-type females the offspring have all been wild-type as expected.

Reviewing the evidence as a whole, therefore, we see that all the possible types of matings have been made and all have given results indicating typical sex-linked inheritance, except in the one generation following the outcross to California stock as discussed above. The absence of swollen females in this generation is presumably due to some genetic difference between the two stocks used, but in any case it appears to bear no relation to the problem under consideration and may be left out of account here.[8]

Since the genetic evidence reveals the type of inheritance expected where the female is XX and the male XY we conclude that sex chromosomes of this type are present[9] and are responsible for sex determination.

Such a conclusion leaves the rôle of the "sex-limited" chromosomes unexplained and also introduces difficulties in any explanation of the "unisexual" progenies, for which the female is responsible. These features will be considered subsequently.

Summary.—See first two and last paragraphs.

[1] This investigation has been aided by a grant from the NATIONAL RESEARCH COUNCIL, Committee for Research on Problems of Sex.

[2] Metz, C. W., *Amer. Nat.*, **40**, 42–56, 1926.

[3] Metz, Moses and Hoppe, *Zeits. ind. Abs. Vererb.*, **42**, 237–270, 1926.

[4] Metz and Moses, *Anat. Rec.*, **34**, 170, 1926. Moses and Metz, these PROCEEDINGS, **14**, 928 (Dec., 1928.)

[5] These have no counterparts in the females, which thus have two chromosomes less than the males.

[6] All our previous material had been secured in the eastern states.

[7] In the males the proportion of swollen flies often exceeds 50%. The reason for this is not yet clear, but in any case it apparently has no bearing on the main question under consideration.

[8] This particular type of cross is being repeated to see if the result is due to a difference between the two strains of flies used. Such a difference might, e.g., result in a selective segregation of chromosomes in the female-producing female analogous to that which has been found regularly in the male. Cf. Metz, C. W., *Zeits. ind. Abs. Vererb.*, **45**, 184–201, 1927.

[9] The presence of Y is inferred from cytological evidence; see papers referred to in footnotes 2 and 3 above.

A SPONTANEOUS MUTATION IN THE HOUSE MOUSE

BY GREGORY PINCUS[1]

BUSSEY INSTITUTION OF HARVARD UNIVERSITY, FOREST HILLS, BOSTON, MASS.

Communicated December 26, 1928

In the 9th brother-sister generation of an inbred line of chinchilla non-agouti (black) piebald mice a single female, No. 7048, was discovered, all of whose ventral hairs (with the exception of a small black area between the forelegs) were white. This ventral lightness was not an extension of piebald spotting, the white being less intense than piebald white, so that a small belly spot of lighter hue due to the piebald factor was quite distinct.

That this animal represented a mutation was apparent for the following reasons: (1) Its parents ♀ 6383 and ♂ 6382 produced in all 14 young in four litters, only one of which, No. 7048, showed the ventral coloration described; (2) ♀ 7048 bred to totally unrelated non-agouti dark-bellied males produced about equal numbers of white-bellied to dark-bellied young (20 light-bellied black: 10 dark-bellied black); (3) three of her sis-

ters backcrossed to her father produced 16 young all dark-bellied; (4) a brother and sister mated together produced 9 young, all dark-bellied. From these facts alone it will be seen that white belly represents a dominant mutation in one of the parental chromosomes so that a single animal heterozygous for the mutant factor appeared.

That the mutant animal could not have arisen by a chance mating of its mother to a black and tan male is evident from the following facts. First, no black and tan animals were in the laboratory at the time the mutant was born, though some were received from Dr. Dunn sometime after. Second, the mutant animal resembled her litter mates in every respect, that is, she carried the recessive genes for piebald and for chinchilla, neither of which occur in any known black-and-tan stock.

Three white-bellied black males derived from the outcross of ♀ 7048 to an unrelated non-agouti stock were used for the production of a test population.

The presence of the chinchilla factor in the mutant animal made it desirable to eliminate this factor because of its known effect in reducing the amount of yellow in the pelage. The animals produced for testing showed white bellies like the original mutant type, but the contour hairs at the edge of the white area were yellow. They were heterozygous for chinchilla (Cc^{ch}).

The close resemblance in appearance of the mutant type to black and tan (Dunn, 1928)[2] made it desirable to discover whether it is like black and tan of allelomorph of the agouti series. It soon became evident that the mutation is an agouti allelomorph and that it is, in fact, black and tan.

In the cross of the mutant type to agoutis (Aa or AA) or to yellows carrying agouti (A^yA) there are produced in F_1 white-bellied agoutis (some of these F_1's showed a cream-colored belly). The F_1 white-bellied agoutis were backcrossed to non-agoutis. If the mutant gene is not an agouti allelomorph then four types should segregate from this backcross: namely, white-bellied agouti, dark-bellied agouti, white-bellied non-agouti, and dark-bellied non-agouti. If, on the other hand, it lies in the agouti series, then only two types are expected in the backcross: dark-bellied agouti and light-bellied black. Similarly from F_1 yellows carrying the mutant gene only two types, yellow and white-bellied non-agouti, are to be expected if the mutant gene lies in the agouti series. Only two types were obtained in each backcross. In the cross of F_1 white-bellied agoutis to non-agouti there were observed 63 agouti to 63 white-bellied non-agouti (expected 63:63). In the cross of F_1 yellows to non-agouti there were observed 19 yellows (or sables) to 18 white-bellied non-agouti (expected 18.5:18.5). The mutant gene was therefore designated a^w. In the course of the experiments a few crosses were made between animals heterozygous for white belly and non-agouti (a^wa) and agouti animals heterozygous for

non-agouti (Aa). Four types would be expected in equal numbers from such a cross:

	Observed	Expected
White bellied agouti (Aa^w)	11	11
Dark bellied agouti (Aa)	11	11
White bellied non-agouti (a^wa)	13	11
Dark bellied non-agouti (aa)	9	11

Tests of dark-bellied agoutis arising in these crosses showed that they did not carry the gene a^w, for on backcrossing them to non-agoutis only two types were obtained, 35 agoutis to 27 non-agoutis (expected 31:31).

It will be seen that these crosses parallel almost exactly those described by Dunn (1928) for black and tan. Some of the mutant type were crossed to black and tans kindly supplied by Dr. Dunn. The F_1 all showed the light belly which was usually somewhat yellowish, but never as intensely yellow as in the pure black and tan. F_2 animals from this cross showed all grades of belly coloration from very white to fairly intense yellow, indicating the segregation of multiple modifiers for yellow pigmentation.

Crosses were also made between white-bellied agouti (A^wA^w) and Dunn's black and tans. As described by Dunn, the F_1's of such a cross are intermediate in belly coloration between yellow and white. The F_2 animals, both agouti and black, show various grades of coloration ranging from white to deep yellow. Backcrosses of F_1's to black and tan gave individuals showing less variation and with more intense yellow coloration than in F_2. A grading system was employed as follows:

Grade 1—white ventral coloration
Grade 2—cream ventral coloration
Grade 3—cream-yellow ventral coloration
Grade 4—yellow ventral coloration.

The results of F_2 and backcrosses to black and tan (B_1) and to my white-belly type (B_2) are summarized in table 1.

Two black females of grade 1 from the F_2 (see table 1) were crossed to ♂ 8112 of my white-belly stock. They produced 14 young graded 1 or 1+. Several F_1 females of the cross white-bellied agouti (A^W) to black and tan (A^T) were mated to non-agouti dark-bellied animals. The grades of the young obtained varied with the male used. When a male (No. 7649) derived from a non-agouti stock showing very intense black pigmentation was used all the offspring (27) were either grade 3 or 4. When a male of a normally intense black was used the grades varied from 1 to 3+ (12 animals).

In short, the yellow-belly coloration of Dunn's black and tan is due, not to the black and tan gene, but to modifiers for intensity of yellow pigmentation. It is of some interest to note in this connection that these

factors making for an intense ventral *yellow* accompany (or are identical with) factors making for intense dorsal black, for the agouti animals of an F_1 population from black and tan show a greater extent and intensity of black on the agouti dorsum than is normally observed.

TABLE 1

GRADES OF VENTRAL COLORATION IN ANIMALS FROM A CROSS OF BLACK AND TAN (A^T) TO WHITE-BELLIED AGOUTI (A^W)

		GRADE 1	GRADE 2	GRADE 3	GRADE 4
F_2.....	Gray	8	13	7	
	Black	3	4	1	1?
	Totals	11	17	8	1
B_1.....	Gray	1?	3	9	6
	Black		4	2	1
	Totals	1	7	11	7
B_2.....	Gray	2	2	2	
	Black	3	1	1	
	Totals	5	3	3	0

Little (1916)[3] has reported exactly parallel results in noting the occurrence of several spontaneous mutations to white-bellied agouti. In one of his experiments a mutant animal arose in a stock which was being selected for increase of extension of black in the agouti pattern. This animal showed a distinctly yellow ventrum comparable to the yellow of Dunn's black and tan. In another experiment where no such selection was practised the three mutant individuals arising showed distinctly whitish ventral coloration.

Whether the light belly of white-bellied agouti and that of my white-bellied black and Dunn's black and tans represent mutation in a locus very closely linked to the agouti locus or whether these are true allelomorphs in the agouti series is a matter of some interest. Its identical dominance in each type and its rather sharp discontinuity in a series of yellow to agouti to black suggests that it may represent a closely linked locus, the white belly appearing in an agouti animal when the mutation occurs on the chromosome containing agouti, and appearing in a black when the mutation occurs on a chromosome containing non-agouti.

It is obvious, in conclusion, that the mutation here reported is the same as the back and tan described by Dunn and represents for all practical purposes a fifth allelomorph in the agouti series in the house mouse.

[1] NATIONAL RESEARCH COUNCIL FELLOW.

[2] Dunn, L. C., *Proc. Nat. Acad. Sci.*, **14**, 816–819, 1928.

[3] Little, C. C., *Amer. Nat.*, **50**, 335–349, 1928.

POSSIBLE α-AUTOMORPHISMS OF NON-ABELIAN GROUPS

BY G. A. MILLER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS

Communicated January 12, 1929

It is well known that a necessary and sufficient condition that a group is abelian is that it admits an automorphism in which every operator corresponds to its inverse. Moreover, if every operator of a group can be made to correspond to its second power, or if every such operator can be made to correspond to its third power in an automorphism it can easily be proved that the group is abelian. In the latter and more difficult case a proof is as follows: Let S_1, S_2 be any two operators of such a group G. It may be assumed that $S_1 \sim S_1^3$, $S_2 \sim S_2^3$, and that $(S_1S_2)^3 = S_1^3 S_2^3$. Hence $(S_2S_1)^2 = S_1^2S_2^2$. Moreover, $(S_2S_1)^3 = S_2^3S_1^3 = S_1^2S_2^3S_1$ and hence $S_2^3S_1^2 = S_1^2S_2^3$. That is, S_1^2 is invariant under G since 3 is prime to the order of G. The relation $(S_2S_1)^2 = S_1^2S_2^2$ therefore implies that $(S_2S_1)^2 = S_2^2S_1^2$, and hence that $S_1S_2 = S_2S_1$.

We proceed to prove that it is possible to find an infinite number of non-abelian groups which admit an α-automorphism, differing from the identity automorphism, whenever α has any given integral value except -1, 0, 1, 2 or 3. It is known that when p is any odd prime number then there is one and only one non-abelian group of order $p^{\alpha+1}$ which involves an operator s order p^α, $\alpha > 1$, and that this group involves p operators of order p which transforms every operator of the group into its $kp^{\alpha-1} + 1$ power, where k has any one of the values 1, 2, .., $p-1$. The central of this group is of order $p^{\alpha-1}$. When $\alpha > 2$ and $p = 2$ there exists also a group of order $2^{\alpha+1}$ which satisfies these conditions. The direct product of such a group and any group which involves only operators whose orders divide $p^{\alpha-1}$ will satisfy the same condition.

Let α represent any positive integer greater than 3 and suppose that $\alpha - 1 = p_1^{\alpha_1}p_2^{\alpha_2} \ldots p_\lambda^{\alpha_\lambda}$, where $p_1, p_2, \ldots, p_\lambda$ are distinct prime numbers. If $\lambda > 1$, at least one of these numbers, say $p_2^{\alpha_2}$, is odd and we may extend any group H of order $(\alpha - 1)p_2$, which involves the cyclic group of order $p_2^{\alpha_2+1}$ as a factor of a direct product, by means of an operator t of order p_2 which transforms every operator of H into its αth power, and hence every operator of the resulting group into the same power. When $\lambda = 1$ and $p_1 = 2$, then α_1 must exceed 1 and hence the desired group can again be constructed. The group obtained by forming the direct product of this group and any group composed of operators whose orders divide $\alpha - 1$ will obviously have the same property. That is, when α is any positive number greater than 3 it is always possible to construct an infinite number of different non-abelian groups which have the property

that each of them has an inner non-identity isomorphism in which every operator corresponds to its αth power.

To prove that such an infinite system of groups exists also when α is any integer less than -1 we may proceed as follows: Let $h = 2 - \alpha$ and assume that $h - 1 = p_1^{\alpha 1} p_2^{\alpha 2}, \ldots, p_\lambda^{\alpha\lambda}$, where the symbols $p_1, p_2, \ldots, p_\lambda$ satisfy the same condition as in the preceding paragraph. The group H is constructed just as before but the operator t transforms every operator of H into its $(p_2 - 1)(h - 1) + 1$ power. If we subtract the order of H from the index of this power there results α. That is, this H has an inner automorphism in which every operator corresponds to its αth power, where α is any negative integer less than -1. This completes a proof of the following theorem: *If α is any given integer with the exception of the five integers, -1, 0, 1, 2, 3, then it is possible to construct an infinite number of distinct non-abelian groups which have the property that every operator corresponds to its αth power in a non-identity inner isomorphism of each one of these groups.*

When every operator of a group G can be made to correspond to its αth power in an automorphism of G then the $(\alpha - 1)$th power of every operator of G is found in its central, and every operator of G can be made to correspond to its $(\alpha - 1)$th power in an isomorphism of G'. When G is non-abelian this isomorphism cannot be an automorphism of G since the group formed by these $\alpha - 1$ powers is abelian. Hence it results that when a non-abelian group admits an α-automorphism it must involve an operator besides the identity whose order divides $\alpha - 1$, and the operators whose orders divide $\alpha - 1$ must constitute an invariant subgroup of the group, which gives rise to an abelian quotient-group. While this subgroup is composed of the operators of G which remain invariant under the α-automorphism in question it is not necessarily composed of invariant operators of G. Moreover, it does not necessarily contain all the invariant operators of G.

As an instance of a group which involves invariant operators which are not invariant under a possible α-automorphism we may cite the non-abelian group of order 64 which contains two cyclic subgroups of order 32. It is obvious that this group admits an α-automorphism in which $\alpha = 5$ and that the operators of order 8 which appear in its cyclic subgroups of order 16 are in its central, but they are not invariant under this α-automorphism. It should be noted that this is an outer α-automorphism. In every inner α-automorphism all the invariant operators of the group obviously correspond to themselves under this automorphism and hence the order of every such operator is a divisor of $\alpha - 1$.

It was proved above that when α does not have one of the five values, -1, 0, 1, 2, 3, then it is always possible to find an infinite number of groups which admit a non-identity inner α-automorphism. It is now easy

to prove that whenever α does not have one of these five values then it is also always possible to find an infinite number of groups which admit an outer α-automorphism. In fact, whenever $\alpha > 3$ it is obviously possible to construct a non-abelian group which involves only operators whose orders divide $\alpha - 1$. By forming the direct product of this group and any group of prime order greater than α there results a group which admits outer α-automorphisms. Whenever $\alpha < -1$ we can construct an infinite number of non-abelian groups composed of operators whose orders divide $2 - \alpha$. The direct product of such a group and a group of prime order greater than $2 - \alpha$ evidently admits outer α-automorphisms. That is *whenever α does not have one of the five values, $-1, 0, 1, 2, 3$, it is possible to find an infinite number of non-abelian groups for any given value, of α such that each of them admits an outer α-automorphism.* It may be added that whenever a group admits an α-automorphism, where $\alpha - 1$ is not divisible by more than two distinct prime numbers, it must be solvable in view of the fact that if the order of a group is not divisible by more than two distinct prime numbers it must be solvable.

[1] J. W. Young, *Trans. Amer. Math. Soc.*, **3**, 186 (1902).

ON CONTINUOUS CURVES HAVING CERTAIN PROPERTIES[1]

BY W. L. AYRES[2]

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF TEXAS

Communicated January 5, 1929

A *continuous curve* is a connected im kleinen continuum (bounded or not). A continuous curve M is *cyclicly connected* if every two points of M lie together on some simple closed curve of M.[3] A subset C of M is a *maximal cyclic curve* of M if C is a cyclicly connected continuous curve and M contains no cyclicly connected continuous curve of which C is a proper subset.[3]

In this note we shall consider continuous curves which have the property that every arc of the continuous curve is a subset of a simple closed curve of the continuous curve. Secondly, we shall consider those in which every arc is a subset of an open curve; and finally, those in which every arc is a subset of a ray. We shall find that these restrictions give much simpler types of continuous curves than might be expected.

THEOREM 1. *In order that every arc of a plane continuous curve M be a subset of a simple closed curve of M it is necessary and sufficient that M be either a simple closed curve or the entire plane.*

Proof.—It is evident that the condition is sufficient. We shall show that

it is also necessary. From the hypothesis it follows at once that M is cyclicly connected. Hence, the boundary of every complementary domain of M is either a simple closed curve or an open curve[4] and no point of M is a cut point of M.[5]

Suppose M has a complementary domain whose boundary is an open curve C. Let P be a point of C and P_1 and P_2 be points, one on each of the two rays of C from P. As P is not a cut point of M, there exists an arc P_1ZP_2 of M that does not contain P.[6] The arc P_1ZP_2 contains a subarc Q_1YQ_2 such that Q_1 lies on one ray of C from P, Q_2 lies on the other, and Q_1YQ_2 has no other point in common with C. Let X be a point on that ray of C from Q_1 which does not contain P. Then the arc PQ_2YQ_1X of $C + Q_1YQ_2$ is a subset of no simple closed curve of M.

Thus M is the entire plane or the boundary of every complementary domain of M is a simple closed curve. Suppose there exists a complementary domain of M and let J be its boundary. If $M \neq J$, let P be a point of $M - J$. Let PA be an arc of M joining P to some point A of J such that $(PA - A)\cdot J = 0$. As A is not a cut point of M there exists an arc PB joining P to some point B of J distinct from A, such that $(PB - B)\cdot J = 0$.[6] Now the set $PA + PB$ contains an arc AXB such that $(AXB)\cdot J = A + B$. Let Y and Z be two points of J separating A and B in J. Then the arc $YAXBZ$ of $J + AXB$ is a subset of no simple closed curve of M. Hence in this case, $M = J$.

THEOREM 2. *In order that every arc of a plane continuous curve M be a subset of an open curve of M it is necessary and sufficient that either M be the entire plane or every point of M be a cut point of M.*

Proof.—The condition is necessary. Suppose M is not the entire plane. Let T be an inversion whose center P does not belong to M. Then $T(M) + P$ is a bounded continuous curve and, from our hypothesis, we have that every two points of $T(M)$ lie together on some simple closed curve of $T(M) + P$ that contains P. Hence $T(M) + P$ is cyclicly connected and the boundary of every complementary domain of $T(M) + P$ is a simple closed curve.[7] Inverting back, it follows that the boundary of every complementary domain of M is either a simple closed curve or an open curve.

(*a*) Suppose there is a complementary domain of M whose boundary is a simple closed curve J. If any component K of $M - J$ has two limit points, P and Q on J, there is an arc PXQ lying in K except for its end points.[8] Let Y and Z be points of J separating P and Q on J. Then the arc $YPXQZ$ is a subset of no open curve of M for either Y or Z lies in a bounded component of $M - PXQ$. Then any component K of $M - J$ has just one limit point on J. From our hypothesis it follows that each set K is unbounded; and thus there are only a finite number of such sets K.[9] Then J contains an arc EFG which contains no one of these limit

points. Let EHG be the other arc of J from E to G. The arc EHG is a subset of an open curve of M and thus EFG must contain a limit point of some component of $M - J$. Hence no complementary domain of M has a simple closed curve as its boundary.

(*b*) By the method used in the second paragraph of the proof of theorem 1, we may show that every point of the boundary of a complementary domain of M is a cut point of M. Suppose there exists a non-cut point P of M. The point P belongs to the boundary of no complementary domain of M, so there exists a simple closed curve J of M containing P in its interior.[10] Let H be the maximal cyclic curve of M containing J. Let D be a complementary domain of H and B be its boundary. Every point of B is a boundary point of some complementary domain of M (but different points may be boundary points of different complementary domains) and thus every point of B is a cut point of M. But no point of M is a cut point of H and thus only a countable number of points of B are cut points of M.[11] Hence the supposition that M contains a non-cut point leads to a contradiction.

The condition is sufficient. This is evident if M is the entire plane. Now suppose every point of M is a cut point of M. Let XY be an arc of M from X to Y. Let M_x be a component of $M - X$ not containing $XY - X$. If M_x is bounded, then $M_x + X$ contains two non-cut points of itself.[12] One of these is not X and this one is also a non-cut point of M. But M contains no non-cut point and thus M_x must be unbounded. Then, as $M_x + X$ is a continuous curve,[9] it contains a ray R_x whose vertex is the point X.[13] Similarly M contains a ray R_y with vertex Y and no other point in common with $XY + R_x$. Then $XY + R_x + R_y$ is an open curve of M containing the arc XY.

For further properties of a continuum every point of which is a cut point, see J. R. Kline, "Closed connected sets which are disconnected by the removal of a finite number of points,"[14] and W. L. Ayres, "On continua which are disconnected by the omission of any point and some related problems."[15]

THEOREM 3. *In order that every arc of a plane continuous curve M be a subset of a ray of M it is necessary and sufficient that either M be the entire plane or every point of M, except possibly one, be a cut point of M.*

The author wishes to express his thanks to Professor J. R. Kline for valuable suggestions which shortened considerably the proofs of the preceding results.

[1] Presented to the American Mathematical Society, November 26, 1927.

[2] NATIONAL RESEARCH FELLOW IN MATHEMATICS.

[3] Whyburn, G. T., these PROCEEDINGS, **13**, 31–38 (1927).

[4] Ayres, W. L., *Bull. Acad. Polonaise*, 127–142 (1928), theorem 8.

[5] Whyburn, G. T., *loc. cit.*, theorem 1.

[6] Moore, R. L., *Math. Zeitsch.*, **15**, 254–260 (1922).
[7] Whyburn, G. T., *loc. cit.*, theorem 10.
[8] Wilder, R. L., *Fund. Math.*, **7**, 340–77 (1925), theorem 1.
[9] Ayres, W. L., *Annals Math.*, Ser. 2, **28**, 501–502 (1927).
[10] Whyburn, G. T., *Trans. Amer. Math. Soc.*, **29**, 369–400 (1927), theorem 7.
[11] Moore, R. L., these PROCEEDINGS, **9**, 101–106 (1923).
[12] Moore, R. L., *Bull. Amer. Math. Soc.*, **29**, 299 (1923).
[13] Knaster B., and Kuratowski, C., *Fund. Math.*, **5**, 26 (1924).
[14] These PROCEEDINGS, **9**, 7–12 (1923).
[15] This paper is not yet published but has been submitted to the *Monatsheften für Mathematik und Physik*. For an abstract, see *Bull. Amer. Math. Soc.*, **34**, 430 (1928).

ON SIMPLE CLOSED CURVES AND OPEN CURVES[1]

BY W. L. AYRES[2]

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF TEXAS

Communicated January 5, 1929

In this note we will give a new characterization of simple closed curves and a necessary and sufficient condition in order that a continuous curve contain an open curve. The first is of interest in itself and the second has some applications in the theory of continuous curves.

THEOREM 1. *If no point of a continuum M in space of n dimensions is a cut point of M but every two points of M may be separated by a subset of M consisting of two points, then M is a simple closed curve.*

Proof.—Since every two points of M may be separated by a finite subset of M, the continuum M is a continuous curve.[3] As M contains no cut point it is cyclicly connected,[4] and contains a simple closed curve J. Let K be any component of $M-J$. Since M has no cut point, K has at least two limit points on J, and there exists an arc PXQ whose end points P and Q belong to J and every other point of PXQ belongs to K. Let PYQ and PZQ be the two arcs of J from P to Q. Then P and Q cannot be separated in M by two points for there are three arcs PXQ, PYQ and PZQ of M from P to Q, no two of which have any other point in common. Hence $M-J$ is vacuous as desired.

LEMMA. *In order that a non-cut point P of a continuous curve M in n-dimensional space lie on a simple closed curve of M it is necessary and sufficient that M not contain a sequence of points $P_1, P_2, P_3, \ldots$ having P as its sequential limit point and such that for every value of i, (1) $M-P_i$ is the sum of two non-vacuous mutually separated sets, M_i and M_{pi}, where M_p is the set containing P, (2) M_i is a subset of M_{i+1}.*

Proof.—The condition is necessary. Suppose the non-cut point P is a point of a simple closed curve J of M. Suppose P is the sequential

limit point of a sequence $P_1, P_2, P_3, \ldots$ satisfying conditions (1) and (2). If $i > j$ and $P_i \neq P_j$, then P_j belongs to M_i. As $J-P_i$ is a connected subset of $M-P_i$ containing P, the set $J-P_i$ belongs to M_{pi} for every i. Suppose some point P_k belongs to the simple closed curve J. As $P_k \neq P$ and P is the sequential limit point of $[P_i]$, there is an integer $m > k$ such that $P_m \neq P_k$. The set $J-P_m$ contains P_k and is a subset of M_{pm}. But since $m > k$ and $P_m \neq P_k$, P_k belongs to M_m. Hence J can contain no point P_i and J belongs to M_{pi} for every i. Now let

$$H = \sum_{i=1}^{\infty} M_i \quad \text{and} \quad K = \prod_{i=1}^{\infty} M_{pi}.$$

If Q is any point of M which does not belong to H, then Q belongs to $P_i + M_{pi}$ for every i. But for any integer k there is an integer $m > k$ such that P_k belongs to M_m. Then Q must belong to M_{pi} for every i and thus to K. Hence $H + K = M$. The set H is an open subset of M and thus contains no limit point of K. Every limit point of H in K is a limit point of $[P_i]$. Then P is the only limit point of H in K, and H and $K-P$ are non-vacuous and mutually separated. Thus P is a cut point of M, contrary to hypothesis.

The condition is sufficient. Suppose P belongs to no simple closed curve of M. Let Q be any other point of M and PQ be an arc of M from P to Q. The point P is a limit point of points of PQ which separate P and Q in M.[5] Let $P_1, P_2, P_3, \ldots$ be a sequence of such points of PQ having P as a sequential limit point and such that the subarc PP_i of PQ contains P_{i+1}. For every i let M_i denote the component of $M-P_i$ containing the subset P_iQ-P_i of PQ, and let M_{pi} be $M-P_i-M_i$. The sets M_i and M_{pi} are mutually separated and for every value of i, M_i is a subset of M_{i+1}. But by hypothesis M contains no such sequence of points $[P_i]$. Thus the condition is sufficient.

THEOREM 2. *In order that an unbounded continuous curve M in n dimensions contain an open curve it is necessary and sufficient that M not contain a sequence of points $P_1, P_2, P_3, \ldots$ such that (1) the sequence $[P_i]$ has no limit point, (2) for every i, $M-P_i$ is the sum of two non-vacuous mutually separated sets M_i^a and M_i^b, (3) for every i, M_i^b is bounded and is a subset of M_{i+1}^b.*

Proof.—The condition is necessary. Suppose M contains an open curve C. If M is the entire space, the condition is satisfied obviously. If not, let T be an inversion whose center P does not belong to M. Then P is a non-cut point of the continuous curve $T(M) + P$ lying on the simple closed curve $T(C) + P$. Now suppose M does contain a sequence $[P_i]$ satisfying (1), (2) and (3) of the theorem. Then $[T(P_i)]$ is a sequence having P as a sequential limit point. From condition (2), $T(M_i^b)$ and $T(M_i^a) + P$ are mutually separated. And from condition (3) we

have that $T(M_i^b)$ is a subset of $T(M_{i+1}^b)$. But by the Lemma, $T(M) + P$ can contain no such sequence $[T(P_i)]$.

The condition is sufficient. If M is the entire space, M contains an open curve. If not, let T be an inversion with center P not belonging to M. If $T(M) + P$ contains a sequence $[P_i]$ satisfying conditions (1) and (2) of the Lemma, then the sequence $[T^{-1}(P_i)]$ satisfies the three conditions of the statement of our theorem. But by hypothesis there is no such sequence $[T^{-1}(P_i)]$. Then the hypothesis of the Lemma is satisfied and P lies on some simple closed curve J of $T(M) + P$. The set $T^{-1}(J-P)$ is an open curve of M.

[1] Presented to the American Mathematical Society, November 26, 1927, as a portion of the preceding paper.

[2] NATIONAL RESEARCH FELLOW IN MATHEMATICS.

[3] G. T. Whyburn and W. L. Ayres, *Bull. Amer. Math. Soc.*, **34**, 349–360 (1928), theorem 2.

[4] G. T. Whyburn, these PROCEEDINGS, **13**, 31–38, 1927. All the theorems of this paper, except theorem 10, are extended to n dimensions in my paper, "Concerning Continuous Curves in Space of n Dimensions," which will appear in the *American Journal of Mathematics*.

[5] W. L. Ayres, these PROCEEDINGS, **13**, 749–754, 1927, theorem 1. This theorem is extended to n dimensions in my paper mentioned in reference 4.

ON A RELATION BETWEEN CONFORMAL AND PROJECTIVE GROUPS IN FUNCTION SPACE

BY I. A. BARNETT

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF CINCINNATI

Communicated December 26, 1928

In an important memoir[1] published in 1871, Felix Klein pointed out an interesting relation between the general conformal group in n-space and the projective group of the sphere in $n + 1$ space. His result may be stated as follows:

The continuous group of projective transformations in $n + 1$ space, which leave invariant the unit sphere in this space, may by a stereographic projection be transformed into the general conformal group in n-space.

It is the purpose of this note to show that there is an analogous relation between the projective and conformal transformations in certain types of function spaces which will be defined presently. Not only is it interesting to note the extension of Klein's result, but the method used in the present paper leads to the consideration of a new class of infinitesimal transformations which are a generalization of those used by Kowalewski.[2]

Let R_x denote the space of continuous functions, so that a point of

this space is given as soon as we specify a definite continuous function $\varphi(x)$ where $0 \leqq x \leqq 1$. Let R'_x be the function space consisting of the totality of points each of which is determined by the pair $(\varphi(x), \sigma)$, where σ is a real continuous variable on the interval $0 \leqq \sigma \leqq 1$.

Consider next the totality of transformations in the function space R'_x which takes every straight line of that space, viz.,

$$\varphi(x) = u(x) + \lambda v(x),$$

$$\sigma = r + \lambda s,$$

into another straight line. All these transformations constitute the projective group of the space R'_x and we may show that the infinitesimal transformations of this group are of the form

$$\begin{cases} \dfrac{d\varphi(x)}{dt} = \alpha(x) + \beta(x)\varphi(x) + \int_0^1 \gamma(x, y)\varphi(y)dy + \sigma g(x) \\ \qquad\qquad + \varphi(x)\left[\int_0^1 \epsilon(y)\varphi(y)dy + e\sigma\right], \quad (1) \\ \dfrac{d\sigma}{dt} = a + b\sigma + \int_0^1 k(y)\varphi(y)dy + \sigma\left[\int_0^1 \epsilon(y)\varphi(y)dy + e\sigma\right], \end{cases}$$

where $\alpha(x)$, $\beta(x)$, $\gamma(x, y)$, $g(x)$, $k(x)$, $e(x)$ are arbitrary continuous functions, and a, b, e are arbitrary real numbers.

We shall also find it necessary to consider the sub-group of (1) which leaves invariant the unit sphere in R'_x,

$$\int_0^1 \varphi^2(y)dy + \sigma^2 = 1. \quad (2)$$

From (2), we see that

$$\int_0^1 \varphi(y)\frac{d\varphi(y)}{dt}dy + \sigma\frac{d\sigma}{dt} = 0, \quad \text{when} \quad \int_0^1 \varphi^2(y)dy + \sigma^2 = 1,$$

and by means of (2) we readily find the following relations:

$$\alpha(x) + \epsilon(x) \equiv 0, \ \beta(x) \equiv 0, \ g(x) + k(x) \equiv 0, \ b = 0,$$
$$a + e = 0, \ \gamma(x,y) + \gamma(y,x) \equiv 0.$$

Thus the infinitesimal transformation of the sub-group sought has the form

$$\begin{cases} \dfrac{d\varphi}{dt} = \alpha(x) + \int_0^1 \gamma(x, y)\varphi(y)dy - \sigma k(x) - \varphi(x)\int_0^1 \alpha(y)\varphi(y)dy + e\sigma\varphi(x), \\ \dfrac{d\sigma}{dt} = -e + \int_0^1 k(y)\varphi(y)dy - \sigma\int_0^1 \alpha(y)\varphi(y)dy + e\sigma^2. \quad (3) \end{cases}$$

By analogy with n-space,[3] we may say that the point $(\varphi(x), \sigma)$ is pro-

jected stereographically with respect to the sphere (2) on the linear spread $\sigma = -1$, if the transformed point $\varphi'(x)$ is given by the formula

$$\varphi'(x) = \frac{2\varphi(x)}{1-\sigma}. \tag{4}$$

Let us now transform (3) stereographically by means of (4). From (4) and (3) we find

$$\frac{d\varphi'(x)}{dt} = 2\left[\frac{-e\varphi(x) + \alpha(x) + \int_0^1 \gamma(x,y)\varphi(y)dy - \sigma k(x)}{1-\sigma} + \frac{\varphi(x)\int_0^1 [k(y) - \alpha(y)]\varphi(y)dy}{(1-\sigma)^2}\right]. \tag{5}$$

Making use of the relations

$$\int_0^1 \varphi'^2(y)dy = 4\left(\frac{1+\sigma}{1-\sigma}\right), \int_0^1 \varphi'^2(y)dy + 4 = \frac{8}{1-\sigma},$$

$$\int_0^1 \varphi'^2(y)dy - 4 = \frac{8\sigma}{1-\sigma},$$

we find that (5) reduces to

$$\frac{d\varphi'}{dt} = \int_0^1 \gamma(x,y)\varphi'(y)dy - e\varphi'(x) + \frac{[\alpha(x) - k(x)]}{4}\int_0^1 \varphi'^2(y)\,dy + k(x) + \frac{\varphi'(x)}{2}\int_0^1 [k(y) - \alpha(y)]\varphi'(y)dy. \tag{6}$$

If now we replace $(k-\alpha)/2$, $-e$, $2\epsilon + \alpha$, φ', γ by ϵ, c, α, f, β, respectively, we see that (6) takes the form

$$\frac{df(x)}{dt} = \alpha(x) + cf(x) + \int_0^1 \beta(x,y)f(y)dy + f(x)\int_0^1 \epsilon(y)f(y)dy - \frac{\epsilon(x)}{2}\int_0^1 f^2(y)dy,$$

which is the infinitesimal transformation of the general conformal group in the space R_x given by Kowalewski.

We have proved the following theorem:

The group of projective transformations in the space R'_x which leave invariant the unit sphere may, by a stereographic projection (4), *be transformed into the general conformal group in R_x.*

[1] Felix Klein, "Ueber Liniengeometrie und metrische Geometrie," *Math. Ann.*, Bd. V, 267 (1871).

[2] G. Kowalewski, "Ueber Funktionraume (II Mitteilung)," *Wien Ber.*, **120**, p. 1435.

[3] Cf. H. Werner, "Bestimmung der grössten Untergruppen derjenigen projectiven Gruppe, welche eine Gleichung zweiten Grades in n Veranderlichen invariant lasst," *Math. Ann.*, Bd. **35**, p. 117 (1890).

ON ELEMENTS OF CONTENT IN METRICAL GEOMETRY

BY FRANCIS D. MURNAGHAN

DEPARTMENT OF MATHEMATICS, THE JOHNS HOPKINS UNIVERSITY

Communicated January 5, 1929

If a spread V_r of r dimensions in a space S_n of n dimensions be given parametrically by equations of the form

$$x^k = x^k(u_1, \ldots u_r); \quad k = 1,2, \ldots n,$$

the orientation of V_r at any point is given by means of the alternating contravariant tensor of rank r

$$d(x^{s_1}, x^{s_2}, \ldots x^{s_r}) \equiv \frac{\partial(x^{s_1}, x^{s_2}, \ldots x^{s_r})}{\partial(u_1, \ldots u_r)} du_1 \ldots du_r.$$

The simple product of this by itself furnishes a contravariant tensor of rank $2r$, namely,

$$X^{s_1, \ldots s_r, t_1, \ldots t_r} \equiv d(x^{s_1}, \ldots x^{s_r})\, d(x^{t_1}, \ldots x^{t_r}),$$

which is alternating in the labels s and t. If, then, we are in possession of a covariant tensor of rank $2r$, $B_{s_1 \ldots s_r, t_1 \ldots t_r}$ say, the inner product $B_{\sigma_1 \ldots \sigma_r \tau_1 \ldots \tau_r} X^{\sigma_1 \ldots \sigma_r, \tau_1 \ldots \tau_r}$ yields at each point of the spread V_r an invariant. On account of the alternating character of $X^{s_1 \ldots s_r, t_1 \ldots t_r}$ we may, without lack of generality, take $B_{s_1 \ldots s_r, t_1 \ldots t_r}$ to be alternating in the labels s and t and may agree that $(\sigma_1 \ldots \sigma_r)$ and $(\tau_1 \ldots \tau_r)$ run over the various groups of r distinct symbols out of the set $1,2, \ldots n$. In other words we shall adopt $\left(\frac{1}{r!}\right)^2 B_{\sigma_1 \ldots \sigma_r, \tau_1 \ldots \tau_r} X^{\sigma_1 \ldots \sigma_r, \tau_1 \ldots \tau_r}$ as the invariant which is to form the subject of discussion. If a change of parameters from $(u_1 \ldots u_r)$ to $(v_1 \ldots v_r)$, say, is made this invariant will be multiplied, as is at once evident from the definition of $X^{s_1 \ldots s_r, t_1 \ldots t_r}$ by the square of $\frac{\partial(u_1 \ldots u_r)}{\partial(v_1 \ldots v_r)}$, and so, if we wish an invariant which will yield, on integration over the spread V_r, a number independent of the parametric representation of V_r, we may choose the square root of the previous invariant, i.e.,

$$\frac{1}{r!} \sqrt{B_{\sigma_1 \ldots \sigma_r, \tau_1 \ldots \tau_r} X^{\sigma_1 \ldots \sigma_r, \tau_1 \ldots \tau_r}}.$$

To be explicit we may associate with each V_2 the number

$$\frac{1}{2!} \int \sqrt{R_{\sigma_1 \sigma_2, \tau_1 \tau_2}\, d(x^{\sigma_1}, x^{\sigma_2})\, d(x^{\tau_1}, x^{\tau_2})}$$

where $R_{pq,lm}$ is the Riemann curvature tensor.

The simplest instance of these general remarks occurs when $r = n$, so that really no summation is involved. The differential invariant is simply $\sqrt{B_{1\ldots n,\,1\ldots n}}\, d(x^1, \ldots x^n)$, where $B_{r_1\ldots r_n,\,s_1\ldots s_n}$ is *any* covariant tensor of rank $2n$ which is alternating in the labels r and s. Such a tensor is furnished by the determinant g of the covariant metrical tensor g_{rs}. In fact, $g = \frac{1}{n!}\,\delta^{\alpha_1,\ldots\alpha_n}_{1\ldots n}\,\delta^{\beta_1,\ldots\beta_n}_{1\ldots n}\,g_{\alpha_1\beta_1}\,g_{\alpha_2\beta_2}\cdots g_{\alpha_n\beta_n}$, where $\delta^{r_1\ldots r_n}_{s_1\ldots s_n}$ is a Kronecker delta.[1] It is clear, then, that the g which occurs in the formula $\sqrt{g}\, d(x^1, \ldots x^n)$ for the element of content of a Riemannian space is really a tensor

$$B_{r_1\ldots r_n,\,s_1\ldots s_n} \equiv \frac{1}{n!}\,\delta^{\alpha_1\ldots\alpha_n}_{r_1\ldots r_n}\,\delta^{\beta_1\ldots\beta_n}_{s_1\ldots s_n}\,g_{\alpha_1\beta_1}\cdots g_{\alpha_n\beta_n}.$$

When $r = n - 1$, we need a covariant tensor of rank $2n - 2$, $B_{s_1\ldots s_{n-1},\,t_1\ldots t_{n-1}}$, say, alternating in the labels s and t and, once this is in our possession, it is easy to introduce a contravariant tensor of rank two. In fact, the fundamental significance of the simple Kronecker delta δ^r_s in tensor analysis is due to the fact that when a transformation of variables from x to $\bar{x}$ is made we have $\frac{\partial x^r}{\partial \bar{x}^\alpha}\,\frac{\partial \bar{x}^\alpha}{\partial x^s} = \delta^r_s$. These equations, n^2 in number, tell us that $\frac{\partial \bar{x}^m}{\partial x^s}$ is the cofactor of $\frac{\partial x^s}{\partial \bar{x}^m}$ in the Jacobian determinant, $j = \frac{\partial(x)}{\partial(\bar{x})}$, divided by j. Hence the equation

$$\bar{B}_{s_1\ldots s_{n-1},\,t_1\ldots t_{n-1}} = {}'B_{\sigma_1\ldots\sigma_{n-1},\,\tau_1\ldots\tau_{n-1}}\,\frac{\partial(x^{\sigma_1}\ldots x^{\sigma_{n-1}})}{\partial(\bar{x}^{s_1}\ldots\bar{x}^{s_{n-1}})}\,\frac{\partial(x^{\tau_1}\ldots x^{\tau_{n-1}})}{\partial(\bar{x}^{t_1}\ldots\bar{x}^{t_{n-1}})}$$

(where the prime indicates that the $\sigma_1 \ldots \sigma_{n-1}$ and $\tau_1 \ldots \tau_{n-1}$ run over the various groups of $n - 1$ out of the n labels, $1, \ldots n$) may be rewritten as

$$\bar{B}_{s_1\ldots s_{n-1},\,t_1\ldots t_{n-1}} = {}'B_{\sigma_1\ldots\sigma_{n-1},\,\tau_1\ldots\tau_{n-1}}\,\frac{\partial \bar{x}^{s_n}}{\partial x^{\sigma_n}}\,\frac{\partial \bar{x}^{t_n}}{\partial x^{\tau_n}}\cdot j^2,$$

where $(\sigma_1 \ldots \sigma_n)$ and $(\tau_1 \ldots \tau_n)$ are rearrangements of the n labels of the same classes as $(s_1 \ldots s_n)$ and $(t_1 \ldots t_n)$, respectively. Combining this with our previous result we see that

$$B^{s_n t_n} \equiv \frac{B_{s_1\ldots s_{n-1},\,t_1\ldots t_{n-1}}}{B_{1\ldots n,\,1\ldots n}}$$

is a contravariant tensor of rank two; it being understood that $(s_1 \ldots s_n)$ and $(t_1 \ldots t_n)$ are arrangements of the same class and that $B_{r_1\ldots r_n,\,s_1\ldots s_n}$ is *any* covariant tensor alternating in the labels r and s. It is in this

way that the contravariant tensor g^{rs} appears in Riemannian geometry; in that instance

$$B_{s_1 \ldots s_{n-1}, t_1 \ldots t_{n-1}} = \frac{1}{n-1!} \delta^{\alpha_1 \ldots \alpha_{n-1}}_{s_1 \ldots s_{n-1}} \delta^{\beta_1 \ldots \beta_{n-1}}_{t_1 \ldots t_{n-1}} g_{\alpha_1 \beta_1} \cdots g_{\alpha_{n-1} \beta_{n-1}}$$

and $B_{1 \ldots n, 1 \ldots n} = g$. The argument is perfectly general. Thus the equations

$$\frac{\partial(x^r, x^s)}{\partial(\bar{x}^\alpha, \bar{x}^\beta)} \frac{\partial(\bar{x}^\alpha, \bar{x}^\beta)}{\partial(x^p, x^q)} = \delta^{rs}_{pq}$$

tell us that $\frac{\partial(\bar{x}^j, \bar{x}^m)}{\partial(x^p, x^q)}$ is the cofactor of $\frac{\partial(x^p, x^q)}{\partial(\bar{x}^j, \bar{x}^m)}$ in the Laplacian development of j divided by j. Hence, if we have a covariant tensor of rank $2n - 4$, $B_{s_1 \ldots s_{n-2}, t_1 \ldots t_{n-2}}$ say, alternating in the labels s and t, we know that

$$B^{s_{n-1} s_n, t_{n-1} t_n} \equiv \frac{B_{s_1 \ldots s_{n-2}, t_1 \ldots t_{n-2}}}{B_{1 \ldots n, 1 \ldots n}}$$

is a contravariant tensor of rank 4 alternating in the labels s and t; and so on.

The interest of this way of looking at these matters is that it indicates how to proceed in more complicated situations. Thus, let us suppose that we are confronted with the well-known suggested extension of Riemannian geometry where the element of arc ds is given by $ds^4 = g_{\alpha\beta\gamma\delta}\, dx^\alpha dx^\beta dx^\gamma dx^\delta$, so that the metrical tensor is a covariant tensor of rank four. The element of content in such a space is $\sqrt[4]{g}\; d(x^1 \ldots x^n)$ where $g = B_{1 \ldots n, 1 \ldots n, 1 \ldots n, 1 \ldots n}$ and $B_{r_1 \ldots r_n, s_1 \ldots s_n, t_1 \ldots t_n, p_1 \ldots p_n}$ is a covariant tensor of rank $4n$ alternating in the labels r, s, t and p. It is defined by the formula

$$B_{r_1 \ldots r_n, s_1 \ldots s_n, t_1 \ldots t_n, p_1 \ldots p_n} = \frac{1}{n!} \delta^{\alpha_1 \ldots \alpha_n}_{r_1 \ldots r_n} \delta^{\beta_1 \ldots \beta_n}_{s_1 \ldots s_n} \delta^{\gamma_1 \ldots \gamma_n}_{t_1 \ldots t_n} \delta^{\delta_1 \ldots \delta_n}_{p_1 \ldots p_n} g_{\alpha_1\beta_1\gamma_1\delta_1} \cdots g_{\alpha_n\beta_n\gamma_n\delta_n}.$$

The contravariant metrical tensor is found by constructing in the same way the alternating tensor of rank $4n - 4$

$$B_{r_1 \ldots r_{n-1}, \ldots p_1 \ldots p_{n-1}} = \frac{1}{n-1!} \delta^{\alpha_1 \ldots \alpha_{n-1}}_{r_1 \ldots r_{n-1}} \cdots\cdots g_{\alpha_1\beta_1\gamma_1\delta_1} \cdots g_{\alpha_{n-1}\beta_{n-1}\gamma_{n-1}\delta_{n-1}} \quad \text{and}$$

writing

$$g^{r_n s_n t_n p_n} = \frac{B_{r_1 \ldots r_{n-1}, \ldots p_1 \ldots p_{n-1}}}{g}.$$

where the arrangements $(r_1 \ldots r_n)$, $(s_1 \ldots s_n)$, $(t_1 \ldots t_n)$, $(p_1 \ldots p_n)$ of the n labels are of the same class. The element of content on a two-way is

$\sqrt[4]{B_{\alpha_1\alpha_2,\beta_1\beta_2,\gamma_1\gamma_2,\delta_1\delta_2}\, d(x^{\alpha_1}, x^{\alpha_2}) d(x^{\beta_1}, x^{\beta_2}) d(x^{\gamma_1}, x^{\gamma_2}) d(x^{\delta_1}, x^{\delta_2})}$ where

$B_{r_1r_2,\, s_1s_2,\, t_1t_2,\, p_1p_2} = \frac{1}{2!}\, \delta^{\alpha_1\alpha_2}_{r_1r_2}\, \delta^{\beta_1\beta_2}_{s_1s_2}\, \delta^{\gamma_1\gamma_2}_{t_1t_2}\, \delta^{\delta_1\delta_2}_{p_1p_2}\, g_{\alpha_1\beta_1\gamma_1\delta_1}\, g_{\alpha_2\beta_2\gamma_2\delta_2}$, and so on for any spread V_r of r dimensions.

[1] O. Veblen, *Invariants of Quadratic Differential Forms*, p. 3.

POTENTIALS OF GENERAL MASSES IN SINGLE AND DOUBLE LAYERS. THE RELATIVE BOUNDARY VALUE PROBLEMS

BY G. C. EVANS AND E. R. C. MILES

DEPARTMENT OF MATHEMATICS, THE RICE INSTITUTE

Communicated January 14, 1929

1. *Introduction.*—The potential due to the most general distribution of finite positive and negative mass deposited in a single layer on a closed surface S may be written in the form

$$v(M) = \int_S \frac{1}{MP}\, d\mu(e_p), \tag{1}$$

where the mass function $\mu(e)$ is a completely additive function of point sets e on S. The most general distribution of mass in a double layer on S yields similarly the potential

$$u(M) = \int_S \frac{\cos\,(MP, n_p)}{MP^2}\, d\nu(e_p), \tag{2}$$

where $\nu(e)$ is likewise a completely additive function; here n_p denotes the direction of the interior normal to S at P.

In terms of these potentials, by means of Stieltjes integral equations, one can solve generalized boundary value problems of the first and second kinds. The first boundary value problem is solved by (2) when the limiting values are given of the quantity $\int u d\omega$, extended over an arbitrary portion ω' of S', which is a surface neighboring S, as S' approaches S. In the second boundary value problem, limiting values of the flux $\int dv/dn\, d\omega$ are similarly given, and the problem is solved in terms of (1). Special cases of these problems are the Dirichlet and Neumann problems, respectively, with boundary values summable on S.

2. *Differential Geometry of S and Its Neighborhood.*—The surface S is a simple closed surface having at every point a tangent plane, whose orientation varies continuously with displacement of the point of tangency

on S. Many properties of the potential functions can be made to depend on the assumption of the existence of a constant Γ such that

$$\int_S \frac{|\cos(MP, n_p)|}{MP^2}\, d\omega_p < \Gamma \tag{3}$$

irrespective of the position of the point M, and

$$\int_S \frac{|\cos(QP, n_q)|}{QP^2}\, d\omega_p < \Gamma \tag{4}$$

irrespective of the position of the point Q on S. In this regard, we have the following lemma.

LEMMA. Let $QP = s$ be the arc length of the curve of section of S made by the plane of n_q and P, and $f(s)$ a positive, continuous, non-decreasing function of s such that $\int_0^s f(s)/s\, ds$ converges. If there is a number δ' such that the inequality

$$|\text{angle}(n_q, n_p)| < f(s) \tag{5}$$

holds for all Q,P on S for which $s < \delta'$, then there is a constant Γ such that (3) and (4) are satisfied.

The function $f(s)$ might be, for instance, of the form Cs^r, $r > 0$. But we shall restrict S in the sequel, except in §6, so that $f(s)$ is proportional to s itself,

$$f(s) = Cs. \tag{5'}$$

The immediate effect of this restriction is to insure that there is a neighborhood of S such that through any point of that neighborhood there passes one and only one normal to S, of length sufficiently small.

Consider a family of surfaces $S' = S'_\tau$, each surface of the family being defined by a normal displacement of the point P of S in amount $n(p)$, $|n(P)| < |\tau|$, with $\lim \tau = 0$. The function $n(P)$ is continuous and non-vanishing, with continuous derivatives with respect to displacement on S. Hence, with τ sufficiently small, there is a one-to-one continuous correspondence of points of S with points of S'; moreover, S' has at every point a tangent plane whose orientation changes continuously with respect to displacement of the point of tangency on S'; and within a suitable uniform neighborhood on S' of a point M on it, the projection on the tangent plane at M of any rectifiable arc s will be a rectifiable arc in the tangent plane of length $\geq s/2$.

A regular closed curve w on S or on S' will be defined as a simple closed curve, composed of a finite number of arcs, each with a continuously turning tangent, the two branches which come together making a not-zero angle with each other. From the remark just made it is evident that such a curve, of sufficiently small diameter and in the neighborhood of a point M, projects into a closed regular curve on the tangent plane.

In order to complete the relation of the family of surfaces S'_τ to S, we assume that there are positive constants C, α, δ'' such that if n_m denotes the normal to S' at a point M on S', and n_q the normal to S at Q on S, the inequality

$$| \text{angle } (n_m, n_q) | < C.MQ^\alpha \tag{6}$$

will be valid whenever $MQ < \delta''$. This requirement is satisfied, for instance, if S' is parallel to S internally or externally, or if S' is the kind of family of surfaces admitted in the calculus of variations.

We shall let S' approach S either internally or externally, and express this relation symbolically by the equations

$$\lim \tau = 0 + \quad , \quad \lim \tau = 0 - \quad ,$$

respectively. We wish to study, under these conditions, various quantities associated with (1) and (2).

3. *Generalized and Iterated Integrals.*—When M is not on S, the integrals (1), (2), are Stieltjes integrals in the ordinary sense. But when the integrand may become infinite, the integrals must be interpreted as generalized or Daniell integrals,[1] and built up by limiting processes. In this sense they exist almost everywhere on S when M becomes a point Q on S. The same applies to the integral

$$\frac{dv(M)}{dn_q} = \int_S \frac{\cos (MP, n_q)}{QP^2} \, d\mu(e_p), \tag{7}$$

the point M being on the normal at Q, when M is allowed to coincide with Q. These functions are summable functions of Q for $M = Q$ on S and we have, with the help of the lemma, the identities

$$\begin{aligned} \int_\omega d\omega_q \int_S \frac{\cos (QP, n_p)}{QP^2} \, d\nu(e_p) &= \int_S d\nu(e_p) \int_\omega \frac{\cos (QP, n_p)}{QP^2} \, d\omega_q \\ \int_\omega d\omega_q \int_S \frac{\cos (QP, n_q)}{QP^2} \, d\mu(e_p) &= \int_S d\mu(e_p) \int_\omega \frac{\cos (QP, n_q)}{QP^2} \, d\omega_q \end{aligned} \tag{8}$$

where ω is the region on S interior to w.[2]

Given a completely additive function of point sets, say $\nu(e)$, there is determined uniquely a bounded additive function of curves $\nu(w)$, with regular discontinuities, which represents the same distribution of mass on S. It is given by the generalized integral

$$\nu(w) = \int_S q(P, w) d\nu(e_p) \tag{9}$$

where $q(P, w)$ is the symmetric surface density at P of the region ω bounded by w. Conversely, given $\nu(w)$, the $\nu(e)$ is uniquely determined. The Stieltjes integrals and the generalized integrals may be written with

respect to the functions of curves $\nu(w)$, $\mu(w)$, as well as the functions of point sets $\nu(e)$, $\mu(e)$.[3]

4. *Statement of Boundary Value Problems.*—Let w be a regular closed curve on S, ω its interior region on S, and let $w' = w'(\tau)$, $\omega' = \omega'(\tau)$ be the corresponding curve and region on $S' = S'_\tau$, and consider the quantities

$$U(\tau, w) = \int_{\omega'} u(M)d\omega_m \quad , \quad V(\tau, w) = \int_{\omega'} \frac{dv(M)}{dn_m} d\omega_m$$
$$U(0\pm, w) = \lim_{\tau = 0\pm} U(\tau, w) \quad , \quad V(0\pm, w) = \lim_{\tau = 0\pm} V(\tau, w) \tag{10}$$

where n_m denotes the normal to S' at M. The function $V(\tau, w)$ represents the total flux across the cap bounded by w', and is of course a function of w' alone, whereas $U(\tau, w)$ depends on S' as well as w'.

THEOREM 1. As S' approaches S entirely from the inside or entirely from the outside of S the following equations hold:

$$U(0\pm, w) = \mp 2\pi\nu(w) + \int_S d\nu(e_p) \int_\omega \frac{\cos(QP, n_p)}{QP^2} d\omega_q \tag{11}$$

$$V(0\pm, w) = \mp 2\pi\mu(w) + \int_S d\mu(e_p) \int_\omega \frac{\cos(QP, n_q)}{QP^2} d\omega_q. \tag{12}$$

The proof of this theorem is obtained by considering separately the contributions to the integrals due to a uniform neighborhood of the generic point P and to the rest of S, respectively. A similar result is obtained also for the function

$$V_1(\tau, w) = \int_{\omega'} \frac{dv(M)}{dn_q} d\omega_m$$

where n_q is the direction of the normal to S at Q, on which M lies.

COROLLARY. The value of the quantity $V_1(0\pm, w)$ is given by the right-hand member of (12).

Let now $F(w)$, $G(w)$ be arbitrary bounded additive functions of regular curves on S, with regular discontinuities. The generalized boundary value problems consist in the determination of potentials of a double layer $u(M)$ and potentials of a single layer $v(M)$ which shall satisfy, respectively, the boundary conditions

$$\frac{1+\lambda}{2\lambda} U(0+, w) - \frac{1-\lambda}{2\lambda} U(0-, w) = F(w) \ , \ \lambda \neq 0, \tag{13}$$

$$\frac{1+\lambda}{2\lambda} V(0-, w) - \frac{1-\lambda}{2\lambda} V(0+, w) = G(w), \ \lambda \neq 0, \tag{14}$$

these being generalized problems, roughly speaking, of the first and second kinds. The usual problems are the degenerate cases of these where the $F(w)$ and $G(w)$ are the indefinite integrals of continuous point functions, instead of being the arbitrary mass functions indicated above.[4]

By combining these equations with the equations (11), (12) above, the following Stieltjes integral equations are obtained:

$$\nu(w) = -\frac{\lambda}{2\pi} F(w) + \frac{\lambda}{2\pi} \int_S d\nu(e_p) \int_\omega \frac{\cos(QP, n_p)}{QP^2} d\omega_q \quad (15)$$

$$\mu(w) = \frac{\lambda}{2\pi} G(w) - \frac{\lambda}{2\pi} \int_S d\mu(e_p) \int_\omega \frac{\cos(QP, n_q)}{QP^2} d\omega_q \quad (16)$$

for the determination of the mass function $\nu(w)$ and $\mu(w)$.

5. *Solutions of the Boundary Value Problems.*—By writing

$$\nu(w) = -\frac{\lambda}{2\pi} F(w) + R(w)$$

the function $R(w)$ turns out to be absolutely continuous, and the equation (15) may be reduced by differentiation to the more usual type of Fredholm equation. A similar treatment applies to (16), and the following theorems may be proved.

THEOREM 2. There is one and only one function $u(M)$, given as a potential of a double layer, for which the first boundary value condition (13) holds, and there is one and only one function, $v(M)$, given as a potential of a single layer, for which the second boundary value condition (14) holds, provided that λ is not a characteristic value of the kernel

$$\lambda K(Q, P) = \frac{\lambda}{2\pi} \frac{\cos(QP, n_p)}{QP^2}.$$

THEOREM 3. The value $\lambda = +1$ corresponds to the interior Dirichlet problem, and to the exterior Neumann problem, and is not a characteristic value. Hence, the solutions of the particular problems corresponding to the conditions

$$U(0+, w) = F(w) \quad , \quad V(0-, w) = G(w)$$

are unique, in the respective classes of functions.

THEOREM 4. The value $\lambda = -1$ is a characteristic value. The exterior Dirichlet problem

$$U(0-, w) = F(w)$$

is therefore solvable by means of a function given as a potential of a double layer if and only if

$$\int_S \varphi_1(P)dF(w_p) = 0,$$

where $\varphi_1(P)$ is a solution of the homogeneous equation with kernel

$$\lambda K(P,Q) = -\frac{\lambda}{2\pi}\frac{\cos(QP, n_q)}{QP^2}, \lambda = -1.$$

The function $\nu(w)$ of (2) is then determined except for an arbitrary additive term of the form $C\omega$, which does not change the value of $u(M)$, M exterior to S.

The interior Neumann problem

$$V(0+, w) = G(w)$$

is solvable by means of a function given as a potential of a single layer if and only if $G(S) = 0$. The function $\mu(w)$ of (1) is then determined except for an arbitrary additive term of the form

$$C\int_\omega \varphi_1(P)d\omega_p;$$

the corresponding $v(M)$ given by (1) contains an arbitrary additive constant, if M is interior to S.

6. *Boundary Values of $u(M)$ and dv/dn.*—By means of the same general methods of analysis we may obtain certain results without specializing S to the extent of assuming $f(s)$ a constant times s. We assume then that $f(s)$ is merely the general function of the lemma of §2. The following theorem, which corresponds to the Fatou theorem in the case of the circle, may then be proved.

THEOREM 5. If Q is a point on S where $\nu(e)$ has a derivative $\nu'(Q) = A$, and M is allowed to approach Q along the normal n_q, then

$$\lim_{M = Q \pm} u(M) = \mp 2\pi A + \int_S \frac{\cos(QP, n_p)}{QP^2} d\nu(e_p), \quad (17)$$

and similarly, if $\mu(e)$ has a derivative $\mu'(Q) = A$,

$$\lim_{M = Q \pm} \frac{dv(M)}{dn_q} = \mp 2\pi A + \int_S \frac{\cos(QP, n_q)}{QP^2} d\mu(e_p). \quad (18)$$

Since a completely additive function of point sets has a derivative (that is, a density) almost everywhere, the equations (17), (18) hold almost everywhere on S.

[1] P. J. Daniell, "A General Form of Integral," *Annals Math.*, **19**, 1918 (279–294).

[2] G. C. Evans, "Fundamental Points of Potential Theory," *Rice Institute Pamph.*, **7**, 1920 (252–329), in particular pp. 257–260.

[3] H. E. Bray and G. C. Evans, "A Class of Functions Harmonic in the Sphere," *Amer. J. Math.*, **49**, 1927 (153–180), in particular pp. 159, 169–171.

[4] The generalized second boundary value problem is treated for plane regions from this point of view by E. R. C. Miles, "Boundary Value Problems for Potentials of a Single Layer," *Trans. Amer. Math. Soc.*, in press.

SUMMARY OF RESULTS AND PROOFS ON FERMAT'S LAST THEOREM (Fourth Paper)

BY H. S. VANDIVER

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF TEXAS

Communicated January 2, 1929

If
$$x^l + y^l + z^l = 0 \tag{1}$$
with x, y and z rational integers prime to each other, $xyz \equiv 0 \pmod{l}$ with l an odd prime, this will be referred to as case II of Fermat's last theorem. Set

$$E_n = \epsilon^R ; R = \sum_{k=0}^{(l-3)/2} s^k r^{-2nk}$$

$$\epsilon = \left(\frac{(1-\zeta^r)(1-\zeta^{-r})}{(1-\zeta)(1-\zeta^{-1})} \right)^{1/2}$$

where $\zeta = e^{2i\pi/l}$, and s represents the substitution (ζ/ζ^r), r being a primitive root of l. Also

$$B_1, B_2, \ldots B_a,$$

will designate the Bernoulli numbers, $B_1 = 1/6$, $B_2 = 1/30$, etc.

In the third paper under the above title I proved a theorem which I repeat here for easy reference:

THEOREM I. *Under the assumptions:*

1. *The second factor of the class number of the field $k(\zeta)$ is prime to l.*

2. *None of the Bernoulli numbers B_{nl}, $n = 1, 2, \ldots, (l-3)/2$ is divisible by l^2, the equation* (1) *is impossible in case II.*

In addition to the above result I have also proved the following theorems:

THEOREM II. *Under the assumptions:*

1. *Only one of the Bernoulli numbers (say, B_n) in the set*

$$B_1, B_2, \ldots, B_{\frac{l-3}{2}} \tag{2}$$

is divisible by l.

2. *The Bernoulli number B_{nl} is not divisible by l^3, the equation* (1) *is impossible in rational integers none zero.*

THEOREM III. *If $l \equiv 1 \pmod 4$ and all the numbers in the set* (2) *which are divisible by l have even subscripts, then* (1) *is impossible in rational integers none zero.*

THEOREM IV. *Under the assumptions:*

1. *None of the units E_a, $a_1 = a_1, a_2, \ldots, a_s$, is congruent to the l-th power of an integer in the field $k(\zeta)$ modulo $\mathfrak{p}$ where $\mathfrak{p}$ is a prime ideal divisor of p; p is a prime $< l^2$, of the form 1 mod l; and $a_1, a_2, \ldots, a_s$ are the subscripts in the Bernoulli numbers in the set* (2) *which are divisible by l.*

2. *The relation*

$$2^{\frac{p-1}{l}} \not\equiv 1 \pmod p$$

holds; the equation (1) *is impossible in case II.*

Of the above theorems the first two are extensions of the results that Kummer essayed to prove in his Memoir of 1857, but the last two are largely new. With the aid of several assistants I have applied the criteria given in each of the four theorems to the exponents in (1) less than 211. The first step was to determine all the non-regular primes less than 211; none were found excepting 37, 59, 67, 101, 103, 131, 149 and 157. Theorem III gives proofs for the exponents 37 and 101 only, but the other three theorems each yielded proofs for all the above non-regular primes mentioned excepting 157, for which case, theorems I and IV yielded proofs. We may then state the following results:

THEOREM V. *The equation*

$$\alpha^l + \beta^l + \gamma^l = 0$$

is impossible for all odd prime exponents l less than 211 *excepting possibly l = 37, 59, 67, 101, 103, 131, 149 and* 157, *if α, β and γ are integers in the field k (ζ), none zero.*

THEOREM VI. *The equation*

$$x^n + y^n + z^n = 0$$

is impossible in rational integers none zero if

$$2 < n < 211.$$

Full details of the proofs of the above results will appear in a paper to be presented to the Transactions of the American Mathematical Society.

THE ANALYSIS OF A CONCRETE CASE OF INTERGRADATION BETWEEN TWO SUBSPECIES[1]

BY FRANCIS B. SUMNER

SCRIPPS INSTITUTION OF OCEANOGRAPHY OF THE UNIVERSITY OF CALIFORNIA

Communicated January 5, 1929

The subdivisions of a species which are known as subspecies or geographic races commonly occupy distinct but adjoining territories, and intergrade with one another along a zone of varying width, lying between these. It may be assumed that two contiguous races of this sort are commonly very closely related to one another, as compared with any two species taken at random. In cases where one of these related forms occupies a much larger and more typical land area than the other, and where it exhibits much more nearly the average characters of the genus, we may reasonably suspect it of being ancestral to its neighbor, particularly if the latter is a highly aberrant form, occupying a restricted and peculiar type of environment.

The intensive study of one concrete example of subspecific intergradation was undertaken by the author in 1927. On a previous trip to the southeastern states during the summer of 1924, I found what appeared to be an extremely promising case for an investigation of this sort.[2] A rather widely distributed mouse, *Peromyscus polionotus polionotus* (Wagner), having color characters fairly typical for the genus, is represented in the coastal region of northwestern Florida and southern Alabama by a quite distinct race (*P. p. albifrons* Osgood), differing from it strikingly in color and some other characters. On an island reef, skirting this part of the Florida coast, is a third race, *leucocephalus* Howell,[3] having paler coloration and more extensive white areas than any other wild mouse with which I am acquainted. This island, as well as the beaches and dunes of the mainland, consist of extremely white sand, covered to only a limited extent by vegetation.

The interior race, *polionotus*, and the coastal race, *albifrons*, although of strikingly different appearance, had been described as subspecies, rather than distinct species, because it was assumed that *somewhere* their ranges met and that intergradation occurred. Up to the fall of 1927, however, a belt of unexplored territory, over a hundred miles in width, intervened between the known ranges of these two forms.

The interest of this case was enhanced by the probable phylogenetic relationship between the two races in question. For the subspecies *polionotus*, in view both of its geographic range and of its distinguishing characters, may be regarded as representing the ancestral stock from which

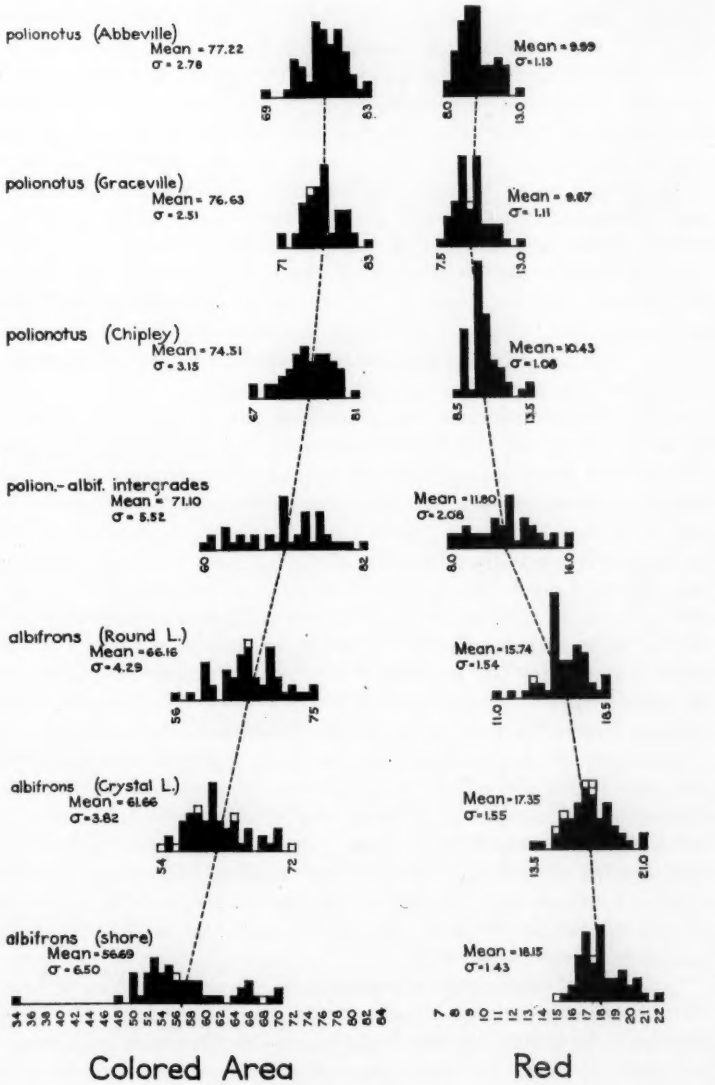


FIGURE 1

Distribution of values for "colored area" and "red," at each of the seven chief collecting stations. The broken lines connect the means of the series. Each square represents an individual. The unshaded squares represent a few individuals, which, for certain reasons, are not included in the regular series, upon which the averages are based.

albifrons immediately sprang. The latter, on the other hand, possesses in a high degree certain peculiarities which have been described in the case of a number of other beach-dwelling types. It seemed possible, therefore, that somewhere in this region the steps might still be discernible through which one geographic race gave rise to another very distinct one.

The present paper is based upon the study of some 400 specimens which were trapped alive by the writer, with the assistance of Mr. J. J. Karol. During September and October, 1927, we trapped at seven chief stations and a few minor ones, situated along a very irregular line, extending from the gulf coast of Florida, near St. Andrews Bay, to the vicinity of Abbeville, in southeastern Alabama. Specimens had previously been collected from the two terminal points of this line by A. H. Howell, but the condition of the intervening population was entirely conjectural.

It was already evident to us, in the field, that specimens trapped close to the gulf shore, were, on the average, paler than those from other points; likewise that there was a gradient of increasing pigmentation as we passed inland, to the northward. Despite these differences, it seemed fair, however, to regard all specimens, up to a point about forty miles from the coast, as being referable to the subspecies *albifrons*. This form possesses a pale, buff pelage, of much the same shade as many of the rodents of our southwestern deserts. Unlike most of the latter, however, the hairs of the ventral white area in *albifrons* are white to their very bases, the dorsal stripe of pigmented hair on the tail is commonly vestigial or lacking, and the dorsal surface of the snout is frequently white. This last is particularly true of the dune-dwelling representatives of the subspecies.

At a point about forty miles from the coast, this pale race was found to give place rather abruptly to one which was far more typical of the genus *Peromyscus*. This was readily recognized as *P. polionotus polionotus*. The latter is a much darker race, quite comparable in appearance with our more familiar white-footed mice of other regions. The dorsal tail stripe is fully developed, and the hair of the ventral area is white at the tips only, being dark at the base. Various regions of the naked skin are likewise more heavily pigmented.

The "area of intergradation" between these two subspecies, at least in the territory explored, was found to be surprisingly narrow. We trapped large and relatively uniform collections of *albifrons* and *polionotus*, respectively, at points not more than ten miles apart, and it is likely that these distances could have been greatly diminished. The population of the intervening area was found to resemble *polionotus* much more nearly than *albifrons*. The majority of specimens appeared to be quite indistinguishable from "true" *polionotus*, although trapped only two or three miles from a region infested by *albifrons*. Other specimens were obviously

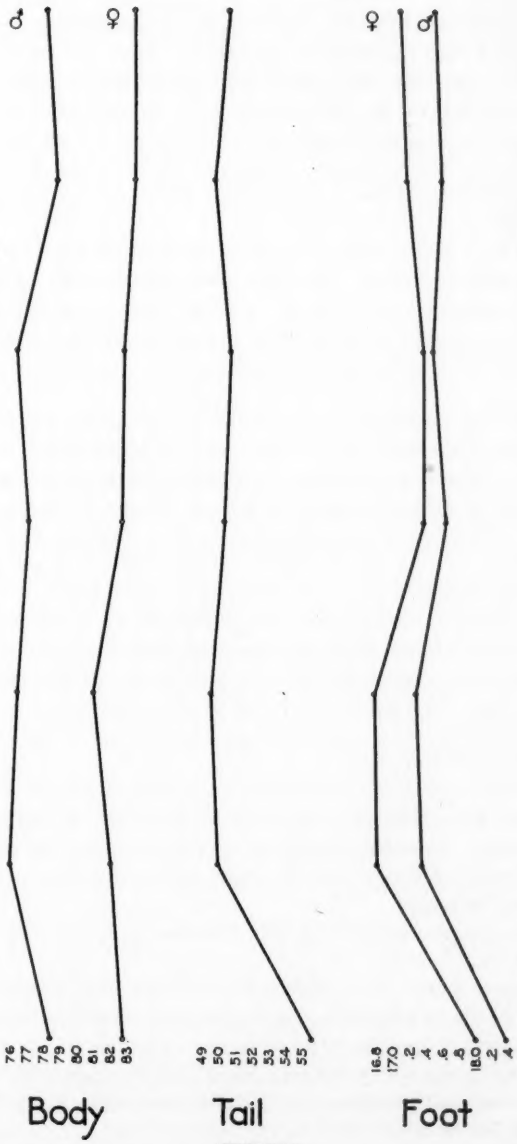


FIGURE 2

Graphs showing the mean values for linear measurements of body (head plus trunk), tail, and foot length, respectively, for each of the seven chief collections. The values for tail and foot have been "corrected" (see text). The seven collecting stations are arranged in the same order as in figure 1.

intermediate between the two races, while a small minority would have passed for the darker specimens of *albifrons*.

This abrupt transition which was found between one subspecies and the other was surprising in view of the lack of any evident barrier, geographic or ecological, between the ranges of the two forms. And it was doubly so when one considered the fact, which had already been determined experimentally, that these forms are perfectly fertile *inter se*, and give rise to fertile hybrids.

The mice from this expedition were nearly all kept for 9 to 11 months in captivity before killing, and were consequently fully mature when they were measured and skinned. The methods employed throughout the *Peromyscus* studies, and the nature of the characters measured have been recently discussed in detail,[4] and need receive only brief mention here.

Tail-length and likewise *foot-length* have here been reduced to a common standard, the values given being those most probable for a body-length of 80 mm. (Computed, of course, by means of the regression coefficient.) "*Colored area of pelage*" represents the percentage of the total pelage occupied by hairs which are pigmented to their tips (thus visible superficially). "*Red*" is the reading obtained through the red color-filter of the Ives Tint Photometer. It is used here merely as an index of the relative paleness or darkness of the hair. "*Pigmentation of ventral hair*" designates the condition of the basal zone of the hairs on this region of the body. Pigment may here be lacking ("0"), or present in full measure ("2"), or only a trace may be present ("1"). The figure for *tail stripe* indicates the length of the dorsal stripe, stated as a percentage of the total length of the visible part of the tail.

Figures 1 and 2 present the distribution frequencies, as well as the means and standard deviations, for five of these characters, in the seven chief local collections, comprising about 50 specimens each. In figure 3 are plotted the mean values of certain pigmental characters, in relation to distance from the coast.[5]

From these graphs, collectively, the following general conclusions may be drawn:

1. In regard to the four pigmental characters here considered, there is a striking degree of parallelism in the geographic relations shown. The condition as regards "red" is by no means an exception, as one might infer, since it is *low* values which indicate *heavy* pigmentation and *high* values which represent *light* pigmentation. To be comparable with the others, this curve should be inverted.

In all cases, there is evident a gradient of varying steepness, from the coast to a point 39 miles in the interior. There is then a very abrupt change within the space of a few miles, the extent of this change, in three

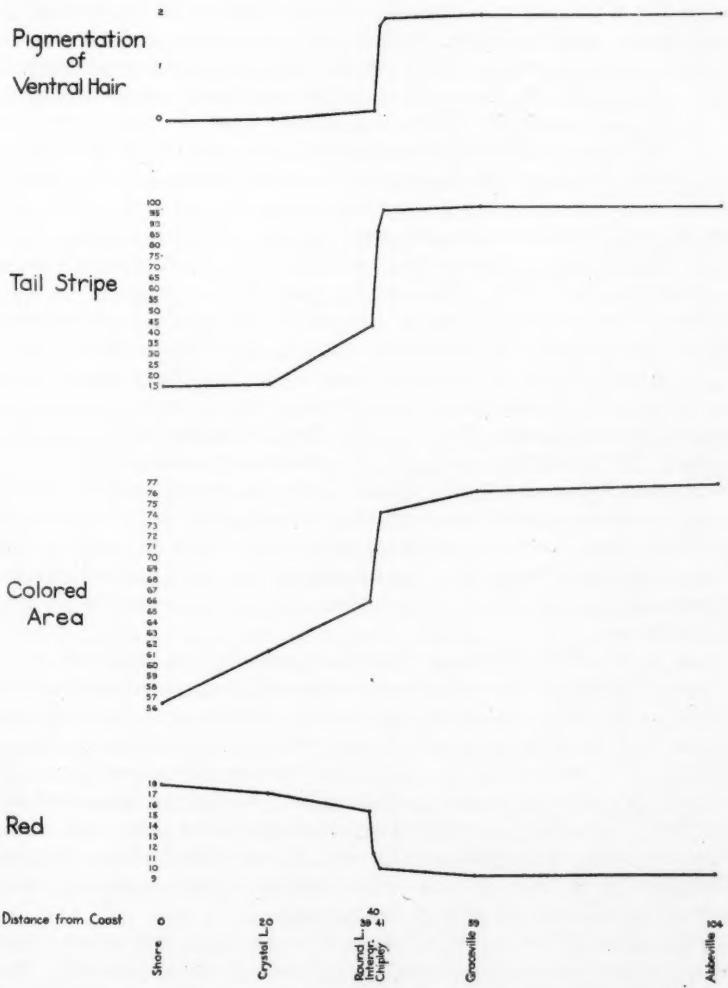


FIGURE 3

Graphs showing mean values for four pigmental characters, in each of the seven collections. In this figure, unlike the others, the stations are plotted with reference to their actual distances from the coast. (Note that these distances represent, in each case, the distance from the *nearest point* on the coast. Since the stations do not lie in a single straight line, this procedure appears to reduce the distance between certain stations.)

cases out of four, being greater than in both of the other segments of the line, taken together. This is followed by a resumption of the original gradient, which, however, is here feeble, and disappears, in all cases but one, between the last two stations. Otherwise stated, *polionotus* displays a much higher constancy within its range than does *albifrons*.

2. Both the graphs for "colored area" and "red," and the standard deviations, computed for these two characters, indicate a considerably lower range of variability for *polionotus* than for *albifrons*. This is in harmony with the last fact mentioned in the preceding paragraph.

3. The group of "intergrades," as might have been expected, shows a higher variability than either of the "pure" races, though its standard deviation for "pigmented area" is exceeded by that of the coast collection of *albifrons* (largely due to a single highly aberrant individual).

4. When we consider the length of tail and foot, a wholly new situation presents itself. It was already known[6] that *albifrons* from coastal points had considerably longer feet and tails than *polionotus*. But the collections of 1927 showed that no such gradient exists as is observed in the case of pigmental characters. The length of both of these members appears to fall off abruptly as we pass to the first inland station, only twenty miles from the shore, and thereafter it undergoes but slight diminution. For both characters, indeed, the minimum length is reached within the range of *albifrons*, a slight initial increase being encountered as we pass to that of *polionotus*.

5. Racial differences in body-length are probably non-significant.

Most students of terrestrial vertebrates will probably agree that the general picture presented by this series of local populations is far from being unique.[7] It is exceptional only in the extent of the divergence shown within a very limited territory, and in the completeness of our knowledge of the case, due to intensive collecting and subsequent analysis of the material. The difficulties which beset an interpretation of this and similar cases, in terms of evolutionary history, are still formidable. I think, however, that we have enough evidence in the present instance to warrant us in making a few tentative suggestions.

Before thus proceeding, it will be well to consider a few facts relative to the genetic status of the characters distinguishing these races. The reader should be reminded that these various racial differences have been definitely proved to be hereditary, at least in certain other species of the genus. The races remain practically unmodified even after a considerable number of generations in a changed environment.[8]

For the present series, I have limited, though fairly decisive, evidence that the most important differences between these local collections are genetic. Finally, even individual differences within a race, in respect to some of these characters, have been shown to be in part hereditary.[9]

Regarding the "behavior" of these subspecific differences in heredity, it may be said that while no single character difference has been detected which acts as if dependent upon a single pair of mendelian factors, I now have abundant and striking evidence of segregation, at least as regards color characters. In even such a wide cross as that between *polionotus* and *leucocephalus*, we find that specimens which reach or surpass the mean value of one or the other parent race, in respect to the extensity of the pigmented area, appear in a quite limited series of F_2 hybrids.

The computation of coefficients of correlation among these various characters, for each of our local collections separately, confirms certain conclusions which have already been reported, for this and related species,[10] but which may once more be profitably summarized. (1) Pigmental characters, as a rule, are at least moderately correlated with one another within the race (mean correlation between "red" and "colored area" is here -0.363). (2) Tail-length and foot-length are moderately correlated, even when the effect of diverse body size has been eliminated. (3) No certain correlation between the length of tail or foot (or of any other body part) and any pigmental character has been discovered.

Data derived both from geographic variation and intra-racial correlation make it plain, therefore, that certain characters commonly vary quite independently of one another, despite instances of simultaneous change along the same environmental gradient. It has likewise been pointed out that the length of these body parts is more subject to erratic and unexplainable local influences and to the transitory effects of environment than are characters related to pigmentation.[11] It seems reasonable, therefore, that we should attach more importance to pigmental characters than to tail or foot-length in considering the probable course of evolution in the present races.

Let us now return to the task of interpreting the relations shown in this cross-section through the ranges of two contiguous subspecies. Firstly, we can hardly escape the conclusion that environment, directly or indirectly, has played some all-important part in the production of these differences. For each of the racial differences here considered appears to depend upon a number of mendelian unit factors, and the process of mutation, unassisted, seems inadequate to account for this piling up of factorial changes *in the same direction.*

Granting this basic fact, we next face the question: How can a continuous environmental gradient bring about such relatively abrupt change in a population as that which is encountered in passing from the range of *polionotus* to that of *albifrons?*

The picture presented in figure 3 suggests rather forcibly that we have to do with one of those critical points in an environmental gradient, at which the individual adjustability of a given race or species seems to reach

its limit. One thinks at once of the analogous situation presented by the familiar "life zones" of our students of distribution. Critical points in the environment do occur, and faunal and floral associations are known to succeed one another in many cases, with considerable abruptness. In the present instance, it would not be unreasonable to suppose that there are fundamental physiological differences, distinguishing these two races, which are quite as pronounced as the visible ones of color and proportions. If one concentrates his attention on these graphs, and abandons himself to an exclusively physico-chemical point of view, it is likely that the foregoing interpretation will assume a high degree of probability. And indeed it has much in its favor. There are, however, serious complications.

In the absence of exact knowledge on this question, the only probable environmental gradients which suggest themselves here are those of temperature and humidity. But the pigmental differences in our present series are not such as we are accustomed to associate with the climatic gradients which here exist. One of the most certain generalizations from distributional studies of mammals and birds is the association of dark pigmentation with high atmospheric humidity. In the present instance, however, our color gradient shows a reverse relation to the customary one, at least, if we may assume a decrease in atmospheric humidity as we pass inland. In any case, the humidity throughout the entire range of *albifrons* probably compares well with that of the ranges of even some of our darker races on the Pacific Coast.

On the other hand, there are strong reasons for believing that extensive depigmentation may result from life upon a background of pale sand. Reference has already been made to the extraordinary pallor of the subspecies *leucocephalus*, in which complete isolation upon a reef of pure white sand, very sparsely clad with vegetation, has entailed an unparalleled degree of modification. Less striking examples have been reported from a number of other beaches where conditions are not so extreme.[12] In the absence of direct evidence, the most plausible hypothesis to account for these changes seems to be that of protective coloration, achieved through the differential survival of paler variants.[13] This is less easy to admit, doubtless, in the case of a nocturnal animal like *Peromyscus*, and it must be borne in mind that the palest of these animals is very far from white. But any appeal to the physical effects of light, direct or reflected, would involve even greater difficulties.

However, we become once more involved in difficulties as soon as we endeavor to apply the protective coloration hypothesis here in detail. It is not obvious why such a highly depigmented condition should occur so far inland, where there are no white beaches or dunes, and where the soil is largely concealed by vegetation. Nevertheless, throughout the entire range of *albifrons* the process of depigmentation has been carried

farther than in any of our desert races of *Peromyscus*, despite the much smaller proportion of bare sand and gravel in the habitat of the former.[14]

Once more, we have left unexplained one of the most striking features in the distribution of these two races, namely, the fact that the boundary between their ranges is so sharp, in spite of their mutual fertility. The simplest explanation of this circumstance would be to assume that our two races became differentiated in localities more remote from one another, and that they subsequently (and rather recently) came into contact. But of this there is no evidence whatever, and there are other circumstances which render such an explanation improbable.

In the present preliminary communication, it is impossible to enter into all the complexities of this situation. A more complete report of these investigations will be published shortly. It must be admitted, however, that the data now at our disposal do not permit us to choose with any confidence among the several alternative interpretations of these facts. As regards the causes which are responsible either for the extent of the respective territories occupied by our two races, or for the abrupt transition between their ranges, we must defer judgment.

However, it can hardly be said that the conclusions herein reached are altogether negative. It has been shown that *albifrons* in all probability arose from *polionotus* (or a common ancestor closely similar to this), in adaptation to life upon a background of extremely pale sand. And it has been possible to portray the probable steps through which the former, highly modified, race arose from the latter. The quality and magnitude of the germinal changes involved in one actual case of evolution are thus to some extent revealed by this analysis, and considerable additional light is thrown upon these changes by the genetic data to be reported shortly.

[1] These studies were conducted under the joint auspices of the Scripps Institution of Oceanography and the Carnegie Institution of Washington.

[2] Sumner, F. B., *J. Mammalogy*, **7**, 1926, 149–184.

[3] Howell, A. H., *Ibid.*, **1**, 1920, 237–240.

[4] Sumner, F. B., *Ibid.*, **8**, 1927, 177–206.

[5] Cf. Cockerell, T. D. A., *Science, N. S.*, **23**, 1906, 145, whose diagram, representing a purely hypothetical case, presents considerable similarity to my own.

[6] *J. Mammalogy*, **7**, 1926, 149–184.

[7] Cf. case of *Neotoma*, discussed by Grinnell and Swarth, *Univ. Calif. Publ. Zoöl.*, **10**, 1913, 197–417 (particularly figs. A,B,C).

[8] Sumner, F. B., *Amer. Nat.*, **58**, 1924, 481–505.

[9] Sumner, F. B., *Ibid.*, **52**, 1918, 290–301. (Also abundant unpublished data.)

[10] Sumner, F. B., *J. Exp. Zoöl.*, **30**, 1920, 369–402; *J. Mammalogy*, **7**, 1926, 149–184.

[11] Sumner, F. B., *Amer. Nat.*, **51**, 1917, 173–185.

[12] Jameson, H. L., *J. Linn. Soc.* (*Zoöl.*), **26**, 1898, 465–473; Bangs, O., *Proc. New Eng. Zoöl. Club*, **4**, 1905, 14–15; Osgood, W. H., *N. Am. Fauna*, U. S. Dept. Agric., **28**, 1909, 285 pp. (particularly 66, 121); Allen, G. M., *Am. Nat.*, **48**, 1914, 467–484; *J. Mammalogy*, **1**, 1920, 235; Sumner, F. B., *Am. Nat.*, **51**, 1917, 173–185; *J. Mammalogy*, **7**, 1926, 149–184.

[13] In admitting the probability that selection, in this case, has operated primarily upon a visible character, color, I am adopting a viewpoint somewhat different from that expressed by me only recently (Carnegie Institution Yearbook for 1928).

[14] Cf. Sumner, F. B., *Ecology*, **6**, 1925, 352–371, where evidence has been presented for believing that the primary cause of the depigmentation of desert animals has been the climatic factor, though the need for concealment (through selection) has probably accelerated the process in certain cases.

NOTE ON C. S. PEIRCE'S EXPERIMENTAL DISCUSSION OF THE LAW OF ERRORS

BY EDWIN B. WILSON AND MARGARET M. HILFERTY

DEPARTMENT OF VITAL STATISTICS, HARVARD SCHOOL OF PUBLIC HEALTH

Read before the Academy November 19, 1928

Incident to the preparation of the official biography of C. S. Peirce for this Academy the senior author came across Peirce's experimental discussion of the law of errors[1] wherein the conclusion was that the normal law was, on the whole, verified. This is in accord with the dictum of Poincaré[2] that everybody believes in the law of errors: the mathematicians because they think it empirically demonstrated by experimenters and the experimenters because they think the mathematicians have proved it *a priori*. The series of observations given by Peirce is long, consisting of about 500 records each day for 24 different days of the time elapsed between the making of a sharp sound and the record of reception of the sound by an observer. According to our previous experience such long series of observations generally reveal marked departures from the normal law, and it seems interesting to examine this material of Peirce's from the point of view of the modern theory of frequency functions.[3] Accordingly we have entered in the tables, (1) the median with its standard deviation, (2) the mean with its standard deviation, (3) the semi-interquartile range, (4) two-thirds of the standard deviation, (5) the ratio of probable error as defined by the semi-interquartile range to the probable error as defined by two-thirds or more accurately 0.6745 of the standard deviation, (6) the arithmetic mean error about the median as origin, (7) the number of negative errors greater than 3σ and the number of positive errors of similar magnitude and their sum, (8) the number of observations within 0.25σ of the mean and the number expected according to the normal law, (9) the percentage excess of the observed over the expected number, (10) the second moment μ_2 and its standard deviation, (11) the third moment μ_3 and its standard deviation, (12) the Pearsonian constant $\sqrt{\beta_1} = \mu_3/\sigma^3$

(Charlier's measure of skewness), (13) the Pearsonian skewness sk, (14) Yule's measure of skewness as the sum of the quartiles less twice the median divided by the semi-interquartile range, (15) the Pearsonian kurtosis β_2-3, (16) the Pearsonian frequency type, with the appended sign + meaning that the sampling error of the fourth moment is infinite and with ++ meaning that the third moment also has an infinite sampling error for samples drawn from the universe defined by the curve.

The following observations may be made on the table:

1. There is on each of the 24 days a greater concentration of small errors in the vicinity of the mean or median than the normal law, fitted with the computed value of the standard deviation, predicts. This may be seen in a number of ways: (*a*) the quotient of the probable error defined in column (3) divided by the probable error in column (4), given in column (5), is uniformly less than 1, varying from a maximum of 0.98 which presumably does not depart significantly from 1 to 0.51 which does greatly depart therefrom, with a mean or median value of about 0.765 which cannot in any way be reconciled with the theoretical value 1. (*b*) The arithmetic mean error in column (6) is never so high as $0.8\,\sigma$ as required by the normal curve but varies from $0.76\,\sigma$ to $0.50\,\sigma$ about a median value of $0.7\,\sigma$. (*c*) The number of observations within $0.25\,\sigma$ of the mean on either side as observed in column (8) is uniformly in excess of the theoretical number by anywhere from 4 to 94%.

2. There is a considerable excess of large errors as compared with the normal law. For example, the normal law would indicate on the basis of 500 observations that one error should exceed $3.1\,\sigma$. In column (7) the observed numbers are entered. There is only one day on which the number of errors above $3.1\,\sigma$ is less than 3 and on five days there are 8. The mean number of negative errors greater than $3.1\,\sigma$ is 1.7 instead of 0.5 and the mean number of positive errors greater than $3.1\,\sigma$ is 3.9 instead of 0.5, making the mean value of the total 5.6 instead of 1. When factors of safety of 3 to 8 are observed, can one say that the theoretical distribution is substantiated? When one calculates the ratio of the largest error to the value of σ for each day one finds numbers running up to over 15; and such errors are inconceivably rare on the basis of the normal law.

As a matter of fact, is it not generally true of errors of observation that they exhibit an excess of small and an excess of large errors as compared with the normal law?

3. The curves exhibit positive skewness. Just how to formulate the precise amount or even the sign of this phenomenon is difficult to say. If the ratio $\beta = \mu_3/\sigma^3$ is taken, column (12), we know that according to the books the value of β for samples drawn from any normal universe is $\sqrt{6/n} = 0.11$ when $n = 500$. There is only one value of β lying between -0.11 and $+0.11$, and the values run up to 10.9 with a median of 1. Two

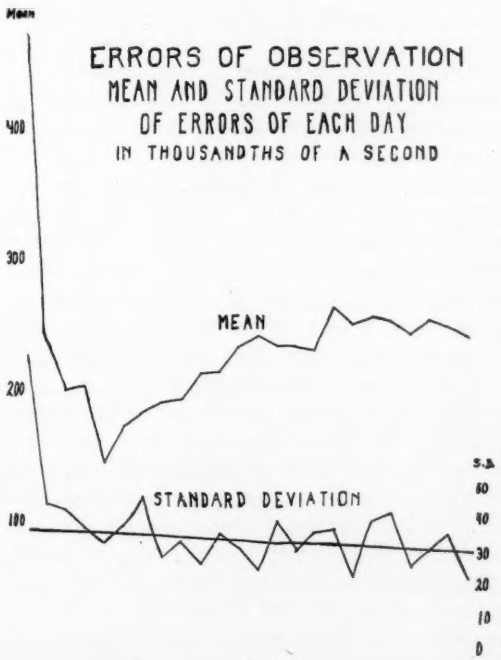
values are negative and 22 are positive. On the basis of Pearson's skewness, column (13), of which the standard deviation is 0.055, there are 4 values less than 0.055 and 20 values greater; there are 10 values less than

	1	2	3	4	5	6	7			8	
DAY	MEDIAN	MEAN	PROBABLE ERROR $\frac{Q_3 - Q_1}{2}$	0.6745σ	$\frac{Q_3 - Q_1}{2(0.6745\sigma)}$	A.D.	ERRORS $>3.1\sigma$ NEG.	POS.	TOTAL	NO. WITHIN 0.25σ OF THE MEAN OBS'D.	EXPECTED
1	468.0±3.0	475.6±4.2	57.9	62.1	0.932	70.1	1	3	4	110	98
2	236.8±2.4	241.5±2.1	26.6	31.6	0.842	35.7	1	0	1	113	97
3	199.9±1.8	203.1±2.0	27.5	30.4	0.905	33.7	0	7	7	113	97
4	201.3±1.2	205.6±1.8	19.4	26.6	0.730	26.9	1	7	8	134	99
5	147.0±2.1	148.5±1.6	21.4	23.5	0.912	26.7	0	4	4	110	97
6	172.4±1.8	175.6±1.8	20.0	26.9	0.744	28.0	0	6	6	119	97
7	184.0±1.6	186.9±2.2	24.6	32.7	0.753	32.3	0	6	6	132	98
8	193.9±1.2	194.1±1.4	17.2	20.5	0.840	22.0	2	6	8	120	97
9	195.4±1.4	195.8±1.6	17.9	23.7	0.756	24.0	2	4	6	132	98
10	215.1±1.5	215.5±1.3	16.4	19.3	0.850	21.0	2	1	3	120	99
11	213.1±1.8	216.6±1.7	19.7	25.2	0.782	25.9	1	5	6	135	99
12	232.7±1.9	235.6±1.7	17.0	22.4	0.759	23.4	3	5	8	103	78
13	244.4±1.3	244.5±1.2	16.3	17.7	0.922	19.9	6	1	7	101	97
14	236.3±1.2	236.7±1.8	14.7	27.8	0.529	21.6	2	3	5	192	99
15	253.3±1.0	236.0±1.4	14.3	21.6	0.662	20.8	4	4	8	162	98
16	233.3±1.4	233.2±1.7	15.6	25.5	0.612	21.7	4	2	6	162	98
17	264.0±1.7	265.5±1.7	20.6	26.0	0.792	27.6	3	5	8	123	100
18	254.0±1.3	253.0±1.1	15.8	16.5	0.959	18.6	0	4	4	114	98
19	255.0±0.9	258.7±1.8	13.9	27.7	0.502	20.4	1	3	4	187	99
20	252.6±1.3	255.4±2.0	15.2	29.2	0.521	21.6	0	3	3	179	98
21	245.2±1.6	245.0±1.2	14.6	18.5	0.790	19.2	3	4	7	120	99
22	255.0±1.5	255.6±1.4	14.8	21.5	0.688	19.0	2	4	6	142	99
23	252.0±1.1	251.4±1.6	15.0	24.6	0.610	19.8	0	3	3	158	98
24	244.2±0.9	243.4±1.1	11.6	15.9	0.730	16.8	3	3	6	113	98
							1.7	3.9	5.6		

	9	10	11	12	13	14	15	16
DAY	EXCESS OBS. OVER EXPECTED %	$\mu_2 \pm \sigma_{\mu_2}$	$\mu_3 \pm \sigma_{\mu_3}$	CHARLIER'S μ_3/σ^3	SKEWNESS PEARSON'S $3k$	YULE'S	$\beta_2 - 3$	PEARSONIAN TYPE
1	12	8475±856	918,117±312,000	1.18	0.417	+0.107	3.1	IV
2	16	2197±168	41,338± 14,000	0.43	0.158	+0.192	0.9	IV
3	16	2046±219	100,824± 37,000	1.09	0.314	−0.029	3.6	IV
4	35	1552±237	111,274± 64,000	1.82	0.407	+0.108	9.7	IV+
5	13	1210± 99	14,511± 11,800	0.39	0.126	+0.252	1.3	IV
6	22	1596±209	91,972± 38,600	1.48	0.375	0	6.4	IV+
7	35	2353±547	336,594±248,000	2.96	0.592	+0.114	24.9	IV++
8	23	926±103	13,856± 12,000	0.48	0.100	+0.070	4.1	IV++
9	35	1233±220	73,081± 59,000	1.71	0.268	−0.056	13.8	IV++
10	21	818±120	18,569± 23,600	0.84	0.129	−0.012	8.8	IV++
11	36	1394±214	83,461± 48,900	1.69	0.352	+0.152	9.8	IV++
12	32	1108±143	22,133± 17,000	0.63	0.118	+0.312	4.7	IV++
13	4	692± 67	4,894± 5,800	−0.22	−0.046	+0.037	2.6	IV+
14	94	1702±616	401,764±207,000	5.74	1.595	−0.034	63.6	VI+
15	65	1030±252	55,666± 25,000	1.68	0.222	−0.042	27.9	IV++
16	65	1426±614	345,877±328,000	6.39	1.445	+0.006	90.6	VI+
17	23	1486±165	14,176± 23,700	0.21	0.039	+0.126	4.3	IV++
18	16	596± 52	3,879± 3,700	0.27	0.080	−0.139	1.8	IV+
19	89	1689±910	758,187±512,000	10.94	112.0	+0.288	143.9	I
20	81	1884±818	632,595±427,000	7.71	3.496	+0.006	91.4	VI
21	21	752±107	4,724± 16,000	0.23	0.030	−0.055	8.2	IV++
22	43	1024±383	171,371±156,000	5.27	1.08	−0.148	68.1	VI++
23	61	1333±343	144,188±135,000	2.73	0.420	−0.220	31.1	IV++
24	15	559± 68	−342± 6,600	−0.02	−0.006	−0.241	5.4	IV++

$3\sigma = 0.165$ but 14 values greater. Still we have two negative and 22 positive values. On the basis of Yule's value, column (14), we find 10 negative, 1 zero and 13 positive values for the skewness; the comparison of the signs and numerical magnitudes in column (14) and in columns (12) and (13) is interesting.

4. The values of the kurtosis are impossible to reconcile with the notion of random sampling from a normal universe. For the value of $\beta_2 - 3$, of column (15), is, as ordinarily given, $0 \pm \sqrt{24/n} = 0 \pm 0.22$, and there is



no value of $\beta_2 - 3$ which is less than 4 times this amount; indeed, the observed values seem to run up almost indefinitely.

The failure of β to lie within the limits ± 0.11 and of $\beta_2 - 3$ to lie within the limits ± 0.22 justifies the statement that on no single day of the 24 was the distribution of errors such that its values of β and $\beta_2 - 3$ could have arisen by random sampling from a normal universe.

5. On account of the large kurtosis the usual formula for the standard deviation of the standard deviation, $\sigma_\sigma = \sigma\sqrt{2n}$, based on the normal curve cannot be applied; we must use $\sigma_\sigma = \sigma\sqrt{(\beta_2 - 1)/4n}$.

6. The third moments as calculated from the data are not reliable

because the sampling errors as calculated from the data, using sixth moments, are so large—see column (11).

7. The type of Pearsonian frequency curve which fits the data is IV on 19 of the 24 days, is VI on 4 days and I on one day. But it should be observed that on 18 of the 24 days the fitted Pearsonian curve has infinite eighth moments μ_8 and that therefore if the universe of errors of observation were really as thus defined on any one of these days the sampling errors of the values of μ_4 would be infinite. The erratic conduct of β_2-3, observed just above, is possibly fair evidence that in very truth the universe of errors of observation in the case of this whole series is heterotypic and that the values of μ_4 or β_2 are in no way to be depended upon. (In 12 of the 24 days the fitted Pearsonian curve has infinite sixth moments so that the reliability of the third moments must be nearly nil—and judging from the sampling errors as given in column (11) it is not alone in these 12 days that the moments are unreliable.)

8. The distributions cannot be represented by the Charlier A-type expansions using the third and fourth moments (but no higher moments) as computed. The departures from the normal curve are in fact so large that the A-type expansion becomes negative in some ranges of the variable less than 3σ removed from the mean in 20 of the 24 days.

The upshot of this all is that Peirce had observations which could show as completely as one might desire that the departures of the errors from the normal law was for his series uniformly great.

Let us turn from the frequency distribution to the statistical constants. The ordinary statement based on the normal law is that the determination of the median is 25% worse than that of the mean. A comparison of the standard deviations of the median and mean in columns (1) and (2) shows that for these observations the median is better determined than the mean on 13 days, worse determined on 9 days, and equally well determined on 2 days. Roughly speaking this means that mean and median are on the whole about equally well determined.

The changes in the mean from day to day are significant in the statistical sense. The value of the mean gives the correction for the "personal equation" of the observer for this type of observation. For scientific purposes in applying the corrections and in estimating the inexactness of the correction we must include in that estimate of inexactness all variations of the personal equation which are not subject to known correction. As the standard deviation of the mean is $\sigma/\sqrt{n}$ or about 1.7 we may not expect better exactness than this. However, the real (as yet uncontrolled and unexplained) variation of the mean is very much greater. Just how to estimate it from the data is impossible to say because there are obviously progressive changes in the mean. On the first day when the observer is "finding himself" the mean is very large; on the second day it drops

to about the terminal value, but continues to drop for three days more; thereafter it rises to a maximum on the 17th day and then falls off.

The actual changes in the mean for the last half of the series are

$$+9, -8, -1, -3, +32, -12, +5, -3, -10, +11, -5, -8.$$

There is little to suggest anything but statistical fluctuations in these differences. It is evident that whatever may be the proper estimate of the standard deviation of the mean even during the last half of the series when the observer may be considered to have settled down, the value must very greatly exceed the 1.7 obtained by the formula $\sigma/\sqrt{n}$, i.e., a reasonably large Lexian ratio is involved. This illustrates the principle that we must have a plurality of samples if we wish to estimate the variability of some statistical quantity, and that reliance on such formula as $\sigma/\sqrt{n}$ is not scientifically satisfactory in practice, even for estimating unreliability of means.

With respect to the variation of the standard deviation from day to day it is clear that the first day is quite imcomparable with the subsequent days, that beginning with the second day there is a gradual reduction of the value of σ of total extent of perhaps 4 or 5 units in 23 days, superposed upon which there are fluctuations of very considerable size. The standard deviation of σ is indeed about 6.2. The formula $\sigma/\sqrt{2n}$ gives about 0.8 for σ_σ on the basis of a mean value $\sigma = 36$. But, as seen above, the kurtosis of the frequency curves is such as to make necessary the use of $\sigma\sqrt{(\beta_2-1)/4n}$ in place of $\sigma/\sqrt{2n}$. The root mean square of $(\beta_2 - 1)/2$ is 3.7 and if it be appropriate to use this as a factor we should give σ_σ the value 3.0 which still is less than half the observed value 6.2.

It appears, therefore, that the sampling theory is not applicable to the results of the identical experiment repeated on 24 days (when the first day is rejected) even when the analysis is confined to the last half of the series. The mean and standard deviation vary much more than is predicted.

[1] On the "Theory of Errors of Observations" by Assistant C. S. Peirce, *Rep. Super. U. S. Coast Survey* (for the year ending Nov. 1, 1870), Washington, Gov't. Printing Office, 1873, Appendix No. 21, pp. 200–224 and Plate No. 27.

[2] H. Poincaré, *Calcul des Probabilités*, chap. X, p. 171, 1912.

[3] The only reference to the series that we have found in the literature is one by M. Fréchet, "Sur la loi des erreurs d'observation," *Rec. Soc. Math., Moscow*, **32**, 1924, in which the author states that at his instigation one of his pupils (Samama) had adjusted by the normal law and by the first Laplacian law $e^{-\alpha|x|}$ several of Peirce's curves and had found that, though the two laws present satisfactory adjustments, the first gives the better result.

PHOTO-ELECTRIC EMISSION AND THERMIONIC EMISSION ONCE MORE

By Edwin H. Hall

Jefferson Physical Laboratory, Harvard University

Communicated December 22, 1928

Doctor L. A. Du Bridge in the *Physical Review* for February, 1928, described experiments which have led him to the conclusion that the photo-electric and the thermionic work functions of out-gassed platinum are equal within the limits of experimental error. He refers also to the work of Warner,[1] who reached a like conclusion with regard to these two work-functions in out-gassed tungsten.

I have no intention of criticizing the experimental accuracy of the work described in either of these cases, but I have a serious, though simple, question to raise concerning the conclusion arrived at by each of the investigators mentioned. Their conclusion will doubtless be generally regarded as meaning that the amount of work required to overcome resisting forces is the same for an electron taken from the "free" state within the metal to the free state outside the metal as for an electron taken by photo-electric action out of the metal. I shall endeavor to show that such an interpretation of the experimental results is not a necessary one.

Let b represent the photo-electric work-function, regarded as independent of temperature. Let us suppose that the thermionic work-function is a somewhat smaller quantity which can be represented with sufficient accuracy as $(b - aT)$, a being a constant and T absolute temperature.[2] In this case the T^2 thermionic emission formula should be written

$$i = AT^2 \epsilon^{-(b-aT)/T} \qquad (1)$$

But, if we let A' be a constant $= A\epsilon^a$, this unfamiliar and therefore repellent equation (1) becomes

$$i = A'T^2\epsilon^{-b/T}. \qquad (2)$$

This differs from the Richardson form only in having an accent on the A, but the b is the constant of the photo-electric work-function.

The experiments of Du Bridge on thermionic emission gave him the value of b, by eq. (2), for out-gassed platinum; and this agreed with the experimentally found photo-electric value. The experiments of Warner did the same with regard to out-gassed tungsten. Neither set of experiments, so it seems to me, throws any light on the value of the constant a in the expression $(b - aT)$ for the necessary work of thermionic emission. Of course, it may be that $a =$ zero, but that remains to be shown.

My belief that a is not zero is based upon studies of conduction and

thermo-electric action made years ago with no reference to the problem of electronic emission.

As to the value of the constant A', usually written A and often called a "universal constant," I predicted[3] a year or two ago that it would prove to be different in different metals. Du Bridge says, regarding this matter, "On computing the value of A from carefully taken emission data on several out-gassed platinum specimens.... values 200 times or more as great as the theoretical were obtained," the "theoretical" value being that suggested by Richardson and Dushman and confirmed, approximately, by Dushman's study of tungsten, tantalum and molybdenum.

[1] *Proc. Nat. Acad. Sci.*, **13**, 56, 1927.

[2] This is not a gratuitous or a new assumption. See my note on the Temperature Relations of Photo-Electric Emission and Thermionic Emission, *Ibid.*, **12**, 486, 1926.

[3] *Ibid.*, **13**, 325, 1927.

THE QUANTUM LAWS AND THE UNCERTAINTY PRINCIPLE OF HEISENBERG

BY GILBERT N. LEWIS AND JOSEPH E. MAYER

CHEMICAL LABORATORY, UNIVERSITY OF CALIFORNIA

Communicated January 11, 1929

In preceding papers[1] we have shown how the second and third laws of thermodynamics may be deduced from a single statistical principle which implied that the properties of a system can be described by assuming a finite number of possible states. This number was called Ω. It is our purpose in this paper to justify and to interpret this assumption, to restate the fundamental laws upon which quantum theory is based, and to show how this restatement leads directly to the uncertainty principle of Heisenberg.[2]

In our previous work, although we endeavored to say nothing that was contrary to the new mechanics, we did not use its phraseology. Indeed, for our statistical purpose, it made no difference whether we considered the number of ways in which certain particles may be distributed among the cells of a phase space, or the number of ways in which the same number of particles with the same energy may be assigned the various discrete solutions of the mechanical equations. The same value is obtained for Ω, for it has been shown by Schrödinger[3] that between any given limits the number of such solutions coincides with the number of cells formerly employed, in the simple cases which have so far been investigated.

Whether this is always true is a question that must be reopened when in the next paper we discuss the electronic states of an atom.

The methods that we have been discussing imply either that every system can occur only in certain definite and fully quantized states, or that a larger number, possibly an infinite continuum of states, is possible; but that within a certain range, that is, within one "cell," these states are for some fundamental reason inherently indistinguishable from one another. These two views seem at first sight to exhaust all the possibilities consistent with a finite Ω. There is, nevertheless, a third view, which is the one we are going to develop here. According to this view, while it is quite permissible to account for the properties of a system by assuming an infinity of states, all distinct from one another, yet it is *sufficient* in order to account for the whole *behavior* of a system to postulate a minimum number of exactly defined states. This minimum number is Ω.

A Restatement of the Quantum Laws.—According to the original law of Bohr, a system in a definite state B with energy ϵ_B changes to the definite state A with energy ϵ_A, emitting, in case the process is one involving light, the energy $h\nu = \epsilon_B - \epsilon_A$. However, we know that this law cannot be exactly true. When allowance is made for Doppler effect and so on, we recognize that there must be left to every spectra line what is known as the natural width. For example, in a gas at very low pressure (as in the nebulæ), there would be no self-reversal of lines if the lines were infinitely sharp, for then no atom would absorb light unless it were exactly at rest with respect to the emitting atom, which it would never be.

It would be natural to ascribe such phenomena as the natural width of the spectral lines to a lack of complete definition of the states, but this is not the only possible explanation. We must consider the possibility that even when two states are exactly defined, the phenomenon occurring in the transition between these states is not uniquely determined.

In much of the physics of this century, there is an element of indeterminism. It is possible that this will later disappear when we take into account for an indefinite time the whole sequence of events in a system, and in all systems with which it is connected. At present we may take the view that when, for example, we know exactly the energy difference between two states, and nothing more, we can say only that the most probable value of the energy given out in the transition is that difference.

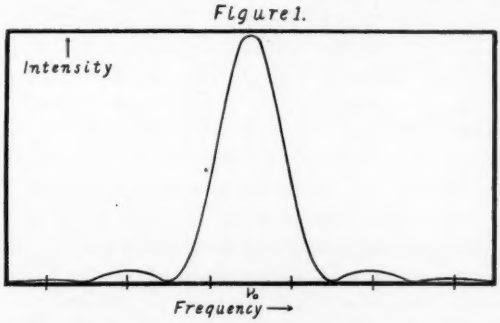
The first step in resolving the indeterminism with which we are here confronted is to ascertain quantitatively the probability that the energy which is lost differs by a certain amount from that most probable value. We shall see that for this purpose we must know, not only the energies of the states, but also their temporal duration. We are familiar with the effect upon the behavior of a system of the proximity in space of other systems; we are less familiar with the effect upon a certain event of the

proximity in time of other events, yet the ideas which we are about to develop are largely those of classical physics.

First, we must assert with all possible emphasis that what we know about the *states* of a system is a deduction from the whole set of phenomena which the system exhibits. In other words, all that we observe is the behavior of a system, after which we invent a set of states which will account in the simplest possible manner for the phenomena observed. For example, we obtain a large number of spectral lines belonging to some gas and then find that these can be interpreted to be the result of transitions between a relatively small number of individual states.

Ordinarily it is advantageous to assume the smallest number of states which will account for the whole behavior, and the search for this minimum number of states is identical with the search for the number Ω,

Figure 1.



which plays so fundamental a part in statistics and thermodynamics. In some cases the particular states which it is necessary to assume are unique. A system with such properties may be called a "definitely quantized" system. In other systems, such as a monatomic gas which has only translational energy, the particular states may be assumed more or less arbitrarily, and only the minimum number of states necessary to account for the observed phenomena of the system is unique. Such a system may be called "arbitrarily quantized."

Returning to the case of the spectral lines, we may regard Bohr's law as an exact limiting law when the several events occurring within a system are infinitely remote from one another in time. In such case, the lines being infinitely sharp, there is no difficulty in ascertaining the states which account for these lines. When, however, the individual states are of short duration the lines are broadened. Can we still explain the phenomena without assuming a larger number of exactly specified states? Our answer, which we shall attempt to justify both qualitatively and quantitatively in the following pages, is that we can.

We shall assume that the states are exact but that the phenomenon observed in a given transition is never twice the same. Thus the energies ϵ_B and ϵ_A of two states are regarded as absolutely fixed, but the energy emitted in the transition will not be equal to $\epsilon_B - \epsilon_A$. It will fluctuate about this value, the average fluctuation being greater the shorter the lives of the individual states. We thus abandon the law of conservation of energy for an individual transition.

Generalizing now from the particular case of spectral lines to all phenomena of all systems, we may postulate the following as the first basic quantum law: I. *All of the phenomena of any system of finite energy may be interpreted by the assumption of a finite number of exactly determined states.*

As a corollary of this law, since the system must always be in one state or another, we may state as the second quantum law: II. *The transition from one state to another is instantaneous.* This corollary is of much significance since, both in the older theories of light and in the newer theories of mechanics, it has often been assumed that an exact mathematical statement concerning an undulatory field would require a knowledge of how the field is built up and how it dies out. According to our view, each field (at a given point) appears instantly, disappears instantly, and remains constant for the time of duration of the given state.

Finally, we may state a third independent quantum law which at present has only an empirical basis, and which we shall not use further in this paper, but is of vital importance to thermodynamics: III. *When the minimum number of states required to account for the behavior of an isolated system has been ascertained, each state will, in the course of time, be found just as often as any other state.*

The Uncertainty in Phenomena Caused by the Finite Life of States.—In classical electromagnetic theory, a system passing, through loss of radiation, from a state B to a state A was believed to generate in space a field, undulating according to the laws of simple harmonic motion. The Poynting vector derived from this field, integrated over a closed surface, gave the rate of loss of radiant energy.

Similarly, the present known facts of optics may be interpreted by assuming that a system in a given state may pass into any one of a number of other states, and that the probability of any one transition, let us say from B to A, is determined by a field, entirely analogous to the electromagnetic field, undulating with a frequency which is exactly equal to $(\epsilon_B - \epsilon_A)/h$. The integrated Poynting vector now gives what is statistically identical with the rate of flow of energy, namely, the probability that the energy associated with this process will be emitted in unit time.

In the electro-magnetic picture, the undulating field might change in amplitude owing to damping, and might perhaps change also in frequency.

Our present picture is a simple one, for we assume that the undulating field is constant both in amplitude and in frequency, from the instant that the state B appears to the instant that it disappears.

In neither picture does the constancy of frequency over a limited period of time imply the emission of monochromatic light.[4] Any simple harmonic motion of limited duration is resolved, either by a Fourier analysis, or by physical apparatus such as a prism or a grating, into a whole continuum of frequencies. The amplitude, at a given distance from the central frequency, is determined by the duration of the simple harmonic motion.

In accordance with the newer trend in mechanics, we may assume that not merely a process in which light is evolved but any process whatsoever is associated with a similar undulating field, which will enable us to predict the probability that the process will occur. Since all of such fields have a temporal variation corresponding to simple harmonic motion, it will be understood that the remarks we are about to make concerning the process of light emission may be applied without essential change to all processes.

Let us assume that, just as in classical theory the squares of the amplitudes corresponding to different frequencies represent the amount of light emitted at these frequencies, so in our present case the squares of these amplitudes represent statistically the amount of energy emitted at the various frequencies when the system is made to drop repeatedly from the state B to the state A.

For simplicity, we shall assume that when a system is in the state B, it has not recently been in that state and will not for a long time appear in that state again. Our mathematical problem then consists in fitting a Fourier integral to a curve whose ordinate is zero from minus infinity to a certain point, follows a simple sine curve for a limited number of periods, and is then zero to plus infinity. The calculation is further simplified by assuming that the number of periods is integral and, though finite, is large. The necessary mathematical steps are given in Mathematical Note I. The result is that if P_ν is the probability that light of frequency ν will be emitted when the system passes from state B to state A,

$$P_\nu = \frac{\sin^2 \pi\tau(\nu - \nu_0)}{\tau\pi^2(\nu - \nu_0)^2}, \qquad (1)$$

where ν_0 is $(\epsilon_B - \epsilon_A)/h$ and τ is the life of the state B.

This equation is plotted in figure 1. It is of the same form as the familiar diffraction curve for a single slit, showing again the essential similarity of time and space. One way to obtain a spectrum with intensities corresponding to figure 1 would be to interrupt a beam of monochromatic light with a shutter operating at a frequency, say, of 10^{-11} or

10^{-12} seconds, when a prism or a grating would give an intensity curve duplicating figure 1.

The General Problem of Uncertainty.—We have seen that by employing methods which are essentially classical we have arrived at the idea that even though the possible states of a system are exactly defined, there is no certainty regarding the phenomena which occur in transitions between these states. Our next step is to show that the uncertainty thus found entirely coincides with the quantitative statement of the uncertainty discovered by Heisenberg.

The *width* of a spectral line may be defined so that the product of this width by the maximum ordinate is equal to the total area under the intensity curve. The area determined by Equation 1 is unity, the maximum ordinate is τ and therefore the width, which we may call $\Delta\nu$ and use as a measure of the uncertainty in frequency, is,

$$\Delta\nu = \frac{1}{\tau}. \qquad (2)$$

It is readily seen that in the present case the width coincides with that distance with respect to which the curve is periodic.

Hitherto we have assumed the life of a given state to be always the same. We shall show, however, (Mathematical Note II) that if the life is variable, following the same law as that of radioactive change, we obtain precisely the same equation for $\Delta\nu$ when $\bar{\tau}$ represents the average life. In such case, τ may be considered as the temporal uncertainty of the process, as $\Delta\nu$ represents the uncertainty in frequency. Representing by $\Delta\epsilon$ the uncertainty in the emitted energy, we may write,

$$\Delta\epsilon = h\Delta\nu, \qquad (3)$$

an equation which according to the new mechanics has a far wider range of validity than the mere processes of light emission. Combining Equations 2 and 3

$$\bar{\tau}\Delta\epsilon = h. \qquad (4)$$

This is one of the important relations of Heisenberg.

We may remark parenthetically, although it is aside from our main topic, that the equation for diffraction from a slit, which corresponds to Equation 1, leads to a similar relation between the width of the diffraction pattern and the width of the slit. Light diffracted by the slit suffers a change in momentum, which is not due to a change in its scalar magnitude, but to a change in direction. The change in momentum corresponding to the width of the diffraction curve we may call Δp and regard as a measure of the uncertainty in momentum due to the slit. The width of the slit l

may be regarded as the uncertainty in position of the light as it passes through. The calculation is entirely similar to the one we have just made, and gives

$$l\Delta p = h. \tag{5}$$

This is the second of the Heisenberg uncertainty relations which, however, we shall presently meet again in a case which has a more direct bearing upon our main problem.

Returning now to figure 1, we shall next examine the case in which a number of such curves overlap. Let us consider, as an example, the familiar case of a molecule which can exist in various vibrational states equally spaced in energy. We may assume that the energies are so nearly alike that the various states occur equally often. For further simplicity, we may ignore rotational levels, and we may also think only of those transitions which occur between the various other states and the zero state. We then have a series of emission lines of equal intensity, equally spaced in frequency, and all having the same width if the various states have the same life. Supposing now, without changing the energy levels, we gradually diminish equally the life of each state (in a similar way, but in a less random way, than would result from an increase in the number of molecular collisions). The emission lines now broaden. A remarkable phenomenon occurs when $1/\tau$ just equals the frequency difference between the spectral lines. Except at the beginning and end of the series, the curves now add exactly to give a horizontal straight line (see Mathematical Note III); in other words, except for the end portions, we obtain a continuous spectrum of constant intensity.

A similar phenomenon is deduced from the diffraction equation. Consider a plane wave of monochromatic light passing in a normal direction through a large number of equal, parallel, long and narrow slits, equally spaced. To one observing in the normal direction from the further side, the slits, if properly spaced, will disappear and seem to be replaced by a field of uniform illumination.

Returning to the vibrational spectrum, it is evident that the whole central portion of this spectrum, which in the particular circumstances that we have just described is entirely uniform, gives no evidence of any particular quantum states. Any section of it, however, can be interpreted by the assumption of a minimum number of quantum states, and this minimum number may be obtained by taking the first state at any arbitrary point and laying off the others at equal intervals, this interval being $1/\tau$. The example that we have given illustrates the sharp transition from the cases in which we determine *what* states must be postulated to account for the behavior of a system, to those cases in which we only determine *how many* states must be postulated. In other words, it is

the transition between systems which are definitely quantized and those which are arbitrarily quantized.

From this example, we obtain a clear picture of the meaning of quantization even in so extreme a case as that of a simple monatomic gas. We may, if we choose, interpret the phenomena exhibited by the gas by assuming an infinite number of perfectly distinct states, but while such a number of states may be sufficient, it is not necessary. Just as in the case of the vibrators, there is a minimum number of discrete states which will explain the observed phenomena, and here again the spacing of these states with respect to energy is determined by the relation

$$\tau\Delta\epsilon = h.$$

In this case, τ, the life of the states, is determined as before by the frequency of events. But these events are now the collisions of the molecules with each other or with the enclosing walls.

As with the vibrators, one of the states may be chosen quite arbitrarily and then the others are laid off at the proper intervals. This produces no ambiguity in determining Ω, except in the immediate neighborhood of the zero state, or the state of lowest energy. Regarding conditions in this small neighborhood, we know nothing at present except that they are of no importance in the great majority of systems with which we have to deal. It remains to show that the minimum number of states which we thus set up is identical with the number of cells into which the phase space is ordinarily divided for statistical purposes, which, in turn, has been shown by Schrödinger to be equal to the number of solutions of his mechanical equations.

Before doing so, lest there be any confusion, we may point out that, in the general case, the states of which we have been speaking are states of the system as a whole. Each event that we have to consider is some transition from one state to another state of the whole system. It sometimes happens, however, and this is true for the ideal vibrator, for the ideal monatomic gas, and for radiation, that we can assign states to the individual particles; and that the properties, such as the energy, of these states are not materially affected by the presence of other particles. In such ideal cases, our procedure is greatly simplified, for if the several atoms are alike and therefore can be assigned the same series of possible states, the minimum number of states Ω for the whole system is merely the number of ways in which all the atoms may be distributed among the several states, in such a way as is consistent with whatsoever conditions are imposed.

In order to show that the minimum number of atomic states which we derive is the same as the number of cells ordinarily used in statistics, let us begin by adopting Ehrenfest's expedient of dealing with a hypothetical

one-dimensional gas. Here we shall say that the life of an atom in a given state is the time that elapses after striking one wall before meeting the opposite wall, assuming that collisions between molecules may be neglected. This is the same for all particles of the same velocity v and if l is the distance between walls,

$$\tau = \frac{l}{v}, \tag{6}$$

but from the fundamental relationship (4),

$$\tau\Delta\epsilon = h,$$

where $\Delta\epsilon$ is the difference of energy between successive states in the particular region of velocity we are considering. This difference in energy requires a difference in momentum Δp. According to the laws of mechanics,

$$\Delta\epsilon = v\Delta p. \tag{7}$$

Eliminating τ and $\Delta\epsilon$ from these three equations, we find finally,

$$l\Delta p = h. \tag{8}$$

If we regard the individual states as defined merely by the momentum, nothing being said about the position of the particle, l may be regarded as the uncertainty in position, while Δp, the difference between the states, is the uncertainty in momentum. Just the same equations are obtained for photons when we consider a one-dimensional Hohlraum.

Proceeding to three dimensions, we may choose arbitrarily three perpendicular directions, x, y and z. We may then define the states of a particle by allowing the momenta p_x, p_y and p_z to assume independently certain discrete values. The problem is to determine the distance apart of these successive values of momentum.

Let us consider two adjacent states in which p_y and p_z are the same, but differing by Δp_x. Designating by $\Delta_x\epsilon$ the difference in energy, we have from mechanics,

$$\Delta_x\epsilon = v_x\Delta p_x, \tag{9}$$

where v_x is, in states of this neighborhood, the velocity in the direction x.

The determination of τ, or what we may here call τ_x, is by no means so obvious as in the one-dimensional case. We may, however, consider the gas as contained in a rectangular box with dimensions l_x, etc., and write in a somewhat formal way,

$$\tau_x = \frac{l_x}{v_x}. \tag{10}$$

Now combining the last two equations with the fundamental Equation 4, we find,

$$l_x \Delta p_x = h. \tag{11}$$

We thus see that our quantization corresponds entirely with the ordinary partition of a phase space into cells.

We shall not at present discuss the more complicated case in which the pressure is so high that the life of a state is determined by collisions between particles rather than by collisions with walls. It should, however, be remarked that the complex phenomenon of the broadening of spectral lines through pressure is probably, in its simplest form, merely the result of shortening the lives of the states through collisions. Hitherto it has been possible to give only a vague definition of a collision, but we may now define it precisely. A collision is any event depending upon the proximity of other molecules which consists in a change of state of one or more molecules. Since the states are usually very close together these collisions occur thousands or millions of times as frequently as the very pronounced collisions of the old kinetic theory. Intermediate between these two are such collisions as have been experimentally studied by Knauer and Stern,[5] where a molecular ray is slightly deflected by gas at low pressure. They found such collisions to occur much more frequently than the collisions of classical theory.

Mathematical Note I.—Starting with a function which has the value

$$f(t') = \alpha \sin 2\pi\nu_0 t' \tag{12}$$

between $t' = -\tau/2$ and $t' = +\tau/2$ (where τ is an integral multiple of the period $1/\nu_0$), and which has the value zero at all other points, this function may be replaced by the Fourier integral

$$f(t) = 2 \int_0^{+\infty} \int_{-\infty}^{+\infty} f(t') \cos 2\pi\nu(t' - t) dt' d\nu, \tag{13}$$

or in differential form,

$$\frac{\partial f(t)}{\partial \nu} = 2 \int_{-\infty}^{+\infty} f(t') \cos 2\pi\nu(t' - t) dt'. \tag{14}$$

The limits of integration may be replaced by $-\tau/2$ and $+\tau/2$, since $f(t')$ is zero outside these limits, and substituting Equation 12, we may perform the integration and obtain

$$\frac{\partial f(t)}{\partial \nu} = \frac{2\alpha\nu \sin \pi\tau(\nu - \nu_0)}{\pi(\nu^2 - \nu_0^2)} \sin 2\pi\nu t. \tag{15}$$

If only those frequencies are significant which are very close to ν_0, as is

the case when τ contains a large number of periods, the equation reduces to the simpler form,

$$\frac{\partial f(t)}{\partial \nu} = \frac{\alpha \sin \pi\tau(\nu - \nu_0)}{\pi(\nu - \nu_0)} \sin 2\pi\nu t. \tag{16}$$

The coefficient which does not involve t may be regarded as a sort of amplitude, a_ν, that is,

$$a_\nu = \frac{\alpha \sin \pi\tau(\nu - \nu_0)}{\pi(\nu - \nu_0)} \tag{17}$$

The energy at each frequency is proportional to a_ν^2, the integral of which over all values of ν proves to be $\alpha^2\tau$. We may therefore normalize the expression for a_ν^2 if we divide by this integral. Thus, if I_ν is the ratio between the energy at a given frequency to the total energy,

$$I_\nu = \frac{\sin^2 \pi\tau(\nu - \nu_0)}{\tau\pi^2(\nu - \nu_0)^2}. \tag{18}$$

Since we have assumed that all the significant frequencies are nearly the same, we may replace I_ν by P_ν where $P_\nu d\nu$ represents the probability that when the transition occurs between state B and state A, the light emitted will have a frequency lying between ν and $\nu + d\nu$.

Mathematical Note II.—It is possible that when we assume the existence of precise quantum states, the life of each state should also be regarded as exactly determined. In that case, it would be necessary to consider such experiments as those of Wien as showing only an *apparent* variability in the life of a state, in the same sense that a study of the energy emitted in a transition shows an apparent variability in the difference of energy between the two states. Nevertheless, we may, for the moment, assume that if we have a large number of identical systems undergoing the same transition, the fraction of these transitions occurring per unit time, $\partial x/\partial t$, follows the simple law of exponential decay,

$$\frac{\partial x}{\partial t} = \frac{1}{\bar{\tau}} e^{-t/\bar{\tau}}, \tag{19}$$

where $\bar{\tau}$ is well known to be the average life. If τ is the actual life in any one system, then all the systems in which the transition occurs at the end of time t have $\tau = t$ and the fraction of these is

$$dx = \frac{1}{\tau} e^{-\tau/\bar{\tau}} d\tau. \tag{20}$$

If we write $P_{\nu,\tau}$ for the characteristic probability of Equation 1 for systems of life τ,

$$P_{\nu,\tau} = \frac{\sin^2 \pi\tau(\nu - \nu_0)}{\tau\pi^2(\nu - \nu_0)^2} \tag{21}$$

Then the characteristic probability at the frequency ν, of all the systems, is obtained by integrating $P_{\nu,\tau}d\tau$. By the last two equations,

$$P_\nu = \int_{\tau=0}^{\tau=\infty} \frac{e^{-\tau/\bar{\tau}} \sin^2 \pi\tau(\nu - \nu_0)}{\bar{\tau}\tau\, \pi^2(\nu - \nu_0)^2}\, d\tau. \tag{22}$$

This is a known integral (see Bierens De Haan, Tables D'Intégrales Définies, Leiden, 1867, Table 365, 4) and,

$$P_\nu = \frac{1}{\bar{\tau}} \frac{\log[1 + 4\pi^2\bar{\tau}^2(\nu - \nu_0)^2]}{4\pi^2\bar{\tau}^2(\nu - \nu_0)^2}. \tag{23}$$

The total integral of P_ν over all values of ν must, of course, be unity, and is so found when the integration is performed (De Haan, Table 134, 12). Since the area under the curve of P_ν is unity the width of the curve $\Delta\nu$, as we have defined it, is the reciprocal of the maximum ordinate. Finding this ordinate to be $\bar{\tau}$ by making $\nu = \nu_0$,

$$\Delta\nu = \frac{1}{\bar{\tau}}, \tag{24}$$

which is the same as Equation 2, except that $\bar{\tau}$ is now the mean life. Equation 23 gives curves which, when spaced at a distance equal to their width, do not, like the curves of Equation 1, add together to give a straight line, but an undulating curve.

Mathematical Note III.—The proof that curves corresponding to Equation 1 add together to a straight line when uniformly spaced at a distance $\Delta\nu = 1/\tau$ is as follows. Consider an abcissa at a distance d from one of the maxima. If n assumes all integral values, the sum of the ordinates at this point will be

$$\sum P = \sum_{n=-\infty}^{n=+\infty} \frac{\tau \sin^2 \pi\tau\left(d + \frac{n}{\tau}\right)}{\left[\pi\tau\left(d + \frac{n}{\tau}\right)\right]^2} = \sum_{n=-\infty}^{n=+\infty} \frac{\tau \sin^2(\pi\tau d + n\pi)}{(\pi\tau d + n\pi)^2} \tag{25}$$

Since the term $n\pi$ contributes nothing to the sine it may be omitted, and removing from the sum the invariant quantities, we find,

$$\sum P = \frac{\tau \sin^2 \pi\tau d}{\pi^2} \sum_{n=-\infty}^{n=+\infty} \frac{1}{(\tau d + n)^2} \tag{26}$$

This sum is known, and is equal to $\pi^2/\sin^2 \pi\tau d$, so that at every point, the sum of the ordinates amounts to τ.

[1] Lewis and Mayer, these PROCEEDINGS, **14**, 569 and 575 (1928).

[2] W. Heisenberg, *Zeits. Physik*, **43**, 172 (1927).

[3] E. Schrödinger, *Phys. Zeits.*, **27**, 95 (1926).

[4] For a discussion of this phenomenon, see W. C. Mandersloot, *Jahrb. Radioakt. und Electronik*, **13**, 1 (1916). Since, however, our assumptions are different from his, our equations are of a different character.

[5] Knauer and Stern, *Zeits. Physik.*, **39**, 764 (1926).

---

*THE TRANSFORMATION SPECTRUM OF THE RUBY**

BY E. L. NICHOLS AND H. L. HOWES

PHYSICAL LABORATORY, CORNELL UNIVERSITY

Read before the Academy November 19, 1928

The luminescence of the ruby under cathodo- and also under photo-excitation has been investigated, among others, by G. C. Schmidt,[1] J. Becquerel,[2] Miethe,[3] DuBois and Elias,[4] Mendenhall and Wood,[5] K. S. Gibson.[6] As the result of their studies the following characteristics have been established:

1. The activator is chromium, presumably in the form of a chromium oxide, in solid solution in aluminium oxide.

2. The sensitiveness of Al_2O_3 to activation by chromium is extreme—so much so that it is difficult to procure or produce Al_2O_3 which does not glow with the characteristic ruddy glow of the ruby when subjected to cathodo-bombardment.

3. Strictly pure Al_2O_3 does not exhibit luminescence of this type.

4. The most striking feature of the luminescence spectrum is the well-known doublet of narrow bands at 0.6918 μ and 0.6932 μ.

5. Coinciding as to wave-length with these are narrow absorption bands and it is possible by suitable adjustments to get reversals suggestive of the reversals of the bright lines of flame spectra.

6. In addition to these bands there are in the spectrum various other narrow bands and that these correspond to absorption regions is strikingly shown by maps of the red portion of the spectrum by DuBois and Elias.

7. The bands of both luminescence and absorption are narrowed to line-like proportions by cooling the crystal under observation to the temperature of liquid air, a fact that was utilized by DuBois and Elias in their search for the Zeeman effect.

The similarity of the phenomena summarized above to those brought out in our own studies of the luminescence of the uranyl salts and of solid

solutions activated by traces of the rare earth elements led us to inquire whether the bands in the spectrum of the ruby occur in sets with constant frequency intervals and further to study the cando-luminescence of this interesting substance. The term cando-luminescence is used here, as in a recent monograph,[7] to denote selective radiation of incandescent bodies not directly ascribable to their temperature. We have found such radiation to be strictly analogous in so many respects to the fluorescence of the same bodies at temperatures below the red heat as to warrant the designation of it as luminescence.

Transformation Bands of the Incandescent Ruby.—To study the visible radiation, a small and somewhat irregular slab of synthetic ruby with unpolished but fairly smooth faces was mounted upon the tip of a vertical splinter of alundum and was heated by the direct application of a hydrogen flame with a reinforcement of oxygen sufficient to give the crystal a temperature somewhat above 1000°C. Under these conditions the spectrum was continuous with regions of unusual intensity in the green and blue. No narrow emission bands were discernible. When the flame was lowered or entirely extinguished and the crystal allowed to cool, various bands appeared, narrow but not line-like. These persisted through a considerable range of temperature, vanishing only as the continuous background disappeared at a lower stage of incandescence. By repeated heatings and coolings it was possible to locate twenty-one such bands between 0.760 μ and 0.426 μ. Some of these could be picked up during rise of temperature also but with greater difficulty.

Temperatures of Appearance and Disappearance.—By the use of a simple device employed and carefully tested out in previous studies of incandescence we were able to determine the upper and lower limits of the range of temperature within which these bands were observed. The method consisted in moistening a minute area, i.e., four or five square millimeters, of the surface of the crystal, with a solution of some uranyl salt, such as the chloride. This upon heating was changed to a thin black coating of oxide which gave with close approximation black-body radiation.

When viewed in the field of an optical pyrometer of the Leeds and Northrup type, the greater brightness of the uncoated crystal as compared with the contiguous "black" patch of uranium oxide was strikingly apparent. This was in agreement with previous studies of the incandescence of the ruby in which three regions of the spectrum (0.42 μ, 0.52 μ and 0.65 μ, respectively) were found to be in excess of ten times as bright at 1000°C. as a black body of the same temperature.[8]

To follow changes of temperature with the pyrometer and to pick up a band promptly as it emerges or disappears, respectively, against the continuous background of the spectrum, requires two observers. Even under

the best conditions the operation is by no means of high precision. A rapid change of temperature favors the prompt recognition of the band but renders the temperature measurement more uncertain. Our determinations gave as the approximate temperature of the appearance of the bands in the red, on cooling 775°C., and for their disappearance about 610°. By rising temperature (a more uncertain observation) 610° for appearance and 750° for disappearance.

We term this banded structure superimposed upon the continuous spectrum the *transformation spectrum* because the definite temperature range within which it is visible coincides with a striking change in the optical properties of the substance. When a ruby is heated it turns black as viewed by reflected light, and the transmitted light loses the ruddy color characteristic of this material when cold. This change is indicative of and is the result of the profound modifications in the general absorption of the crystal described several years ago in a detailed study by K. S. Gibson.[9]

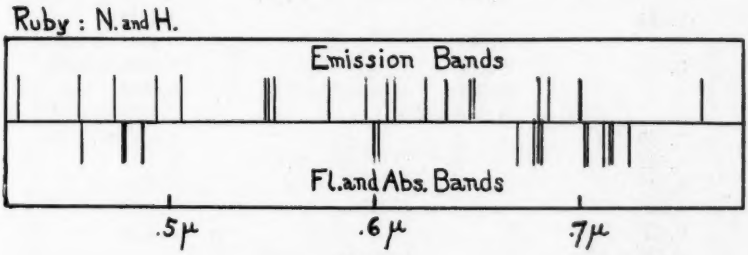


FIGURE 1

The blackening of the ruby upon heating, with loss of color of the transmitted light, was fairly abrupt, and we were able to note that it coincided in temperature with the appearance of the transition bands. A setting with the pyrometer of this change of color in fact gave again the temperature as 610.°

The relation of the twenty-one transition bands to the bands of fluorescence and absorption mapped by DuBois and Elias is somewhat obscured by the very large shift of the bands recorded by these writers toward the longer wave-lengths. That at least one such band persists over a range of more than one thousand degrees (i.e., from −190°C. to + 845°C.) was noted by them and they determined its position at nine temperatures throughout that range, finding a gradual and constant movement from 0.6925 to 0.7033 or 112Å. Their settings were made upon the middle position of the strong doublet R_1R_2 which at higher temperatures tends to merge into a single band. We assumed for the purposes of comparison that all bands of the absorption-fluorescence system undergo a similar shift and adjusted the readings of the twenty-three bands given by them

to positions corresponding to the temperature at which our transformation bands were observed.

As may be seen from Table 1 and also from figure 1 there is no obvious relation between the two, beyond the fact that the emission bands tend

TABLE 1

TRANSFORMATION N. AND H.	ABSORPTION AND FLUORESCENCE DUBOIS AND ELIAS	TRANSFORMATION N. AND H.	ABSORPTION AND FLUORESCENCE DUBOIS AND ELIAS
0.760 μ			0.599–0.602
	0.724 μ		0.600–0.601
	0.716	0.592	
	0.725		
	0.712	0.578	
	0.704	0.552	
	0.703	0.549	
0.700		0.547	
	0.689	0.506	
	0.686	0.494	
0.685			0.4873
	0.680		0.4865–0.4871
0.679	0.679		0.4849–0.4857
	0.678		0.4856
	0.669–0.670		0.4785–0.4794
0.649			0.4777–0.4779
0.647		0.473	
0.640			
0.635			0.4575
0.625			0.4570
0.609–0.610		0.456	
0.607		0.426	

to fill in the gaps between the groups of fluorescence-absorption. This is clearly a result of the greatly increased absorption at the higher temperatures which tends to merge individual bands into regions of continuous absorption and to bring out such bands where none were strong enough to be observed at room temperature.

It does not follow that because among the bands of fluorescence and absorption, corrected for the temperature shift as above described, and the emission bands of the incandescent ruby only one coincidence was observed, the two do not belong to the same system. On the contrary when wave-numbers ($1/\mu \times 10^3$) are taken in place of wave-lengths both fall equally well into the same sets of equal frequency interval, as shown in Table 2. The common frequency difference is 36.7, a value very close to the lower interval given by Tanaka[10] in his study of chromium as an activator when in solid solution in CaO, i.e., 36.4.

A convenient method of showing graphically the relation between sets is to plot wave-numbers ($1/\mu \times 10^3$) as abscissae and to divide the or-

TABLE 2

RELATIVE POSITION OF SETS (INT. 36.7), RUBY SPECTRUM

READINGS N	READINGS D AND E	VARIATIONS	SETS WITH THEIR POSITIONS FOR OCCUPIED STEPS A 1306.5	B 1310.5	B′ 1313.5	C 1316.0	C′ 1318.5	D 1322.8	D′ 1326.2	E 1333.5	STEP
1316		=				1316.0					0
	1381.2	−1.3=	1379.9								2
	1396.6	−0.4=						1396.2			2
	1399.0	+0.6=							1399.6		2
	1405.0	+1.9=								1406.9	2
	1420.1 R	+0.5=		1420.6							3
	1422.9 R	+0.5=			1423.6						3
1428		+0.6=					1428.6				3
	1451.4	+1.9=	1453.3								4
	1457.7	−0.4=		1457.3							4
1460		+0.3=			1460.3						4
1471	1470.8	−1.4, −1.2=						1469.6			4
	1472.8	+0.2=							1473.0		4
	1494 (92.5)	(−1.5)=		1494.0							5
1541		+2.0=						1543.0			6
1546–8		−0.4, +1.6=							1546.4		6
1563		+0.4=	1563.4								7
1575		+0.4=					1575.4				7
1600		+0.1=	1600.1								8
1639–42		+1.8, −1.2=		1640.8							9
1647		−0.7=				1646.3					9
	1663.9[1]	−0.1=								1663.8	9
1678		−0.5=		1677.5							10
1730		−0.1=							1729.9		11
1812		−1.4=								1810.6	13
1821		−0.7=	1820.3								14
1828		−0.7=			1827.3						14
1976		+0.6=				1976.6					18
2024		−0.5=							2023.5		19
	2052.1	+0.4=					2052.5				20
	2053–5.5	+1.3=						2056.8			20
	2059.3[2]	+0.9=							2060.2		20
	2085.9–9.9	+0.8, −0.7=				2086.7	2089.2				21
	2092.5–3.3	+1.0, +0.2=						2093.5			21
2114		−0.1=	2113.9								22
	2185.8 (8.0)	+1.5, −0.7=	2187.3								24
2193		+1.3=			2194.3						24
2347		−0.9=					2346.1				28

[1] Band measured as 1660.6–9.2 and as 1663.9–7.5

[2] Band also measured as 2058.9–62.3.

dinate axes into equal steps each representing the frequency interval. Since the distance between members of a given set is always an exact multiple of the interval the members will occupy the same position within the steps in which they occur and will lie along a vertical line, and the lines representing the various sets will be parallel.

To bring out upon a greatly exaggerated scale the departures from exactitude of the individual readings a form of diagram suggested by Mr. D. T. Wilber and worked out in detail for the data in the present paper by Miss M. A. Ewer may be used. In this form of diagram the entire base line represents the magnitude of a single step and ordinates are plotted as

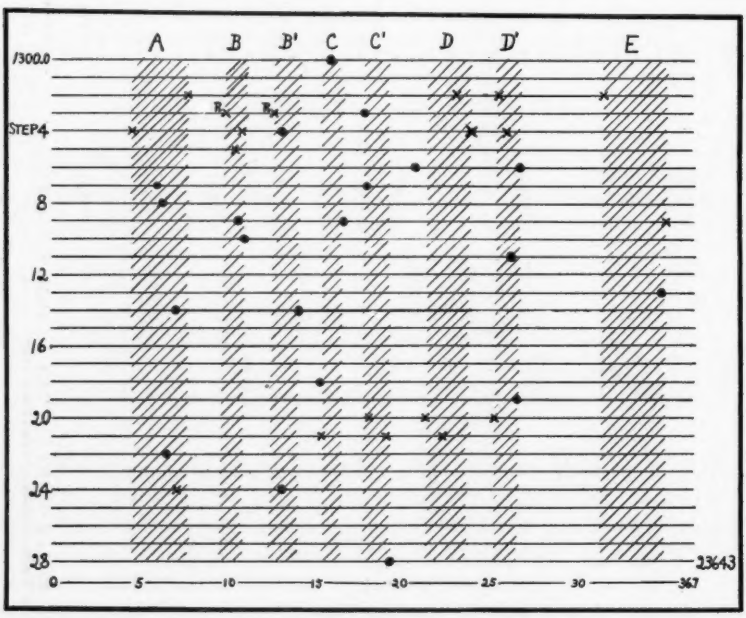


FIGURE 2

wave-numbers arranged in steps. If now readings are expressed in terms of their proper position in the steps in which they occur, by subtracting from each the wave-number corresponding to the beginning of the step,[11] their departures from a vertical line, as in figure 2, will show the size and sign of the errors. In the diagram the vertical shaded areas show for each set the limits within which lie all but one somewhat aberrant reading (representing an emission band belonging presumably to Set D).

The absorption spectrum, in addition to the well-known doublet R_1R_2, with a frequency difference of three units, has other similar pairs of bands;

and if we assume three close-lying pairs of sets with the frequency interval 36.7 (*B-B'*, *C-C'* and *D-D'*), so located as to contain the members of these doublets, also a single set A, all the known bands of fluorescence, absorption, or emission fall into one or another of these sets excepting two absorption bands and one emission band. The three readings thus excluded are so situated that they could not possibly be considered as belonging to any of the above sets. They are separated, however, from each other by distances which are very closely exact multiples of the frequency interval (36.7), and are therefore put in a set by themselves (*E*), of which they may be regarded as the only members to be found under the conditions which happened to prevail. It is obvious that throughout broad regions such bands might be invisible because of too great transparency of the crystal at low temperatures, and be obscured at temperatures of incandescence by the strong continuous radiation.

It is notable that although the transformation bands detected by us and some of the absorption and fluorescence bands recorded by DuBois and Elias were of considerable width, there is no instance where the set to which a band belongs is in doubt. The average departure of the readings from the exact corresponding place in the set is, in fact, only 0.7 and there are only three cases out of forty-four where the departure exceeds 1.6 frequency units.

Summary.—1. The spectrum of an incandescent ruby heated in a hydrogen flame was found to have superimposed upon the general radiation a number of rather narrow emission bands.

2. The temperature-range within which the bands are visible, i.e., circa 600°C. to 775°C., corresponds to profound changes in the absorption of light by the ruby.

3. These so-called transformation bands occur in sets having a constant frequency interval of 36.7, of which sets all the known absorption and fluorescence bands of the ruby are likewise members.

4. The phenomena described in this paper are clearly analogous to those previously observed in the incandescence of solids containing various rare earths as activators.

* An investigation carried out under a grant from the Hecksher Fund of Cornell University.

[1] G. C. Schmidt, *Ann. Physik*, **xiii**, 625, 1904.

[2] J. Becquerel, *Phys. Zeits*, **viii**, 932, 1907.

[3] Miethe, *Verh. d. physik. Ges.*, **xi**, 715, 1907.

[4] DuBois and Elias, *Ann. Physik*, **xxvii**, 233, 1908; **xxxv**, 617, 1911.

[5] Mendenhall and Wood, *Phil. Mag.*, [6], **xxx**, 316, 1915.

[6] Gibson, K. S., *Phys. Rev.*, [2], **viii**, 38, 1916.

[7] Nichols, Howes, and Wilber, "Cathodo-luminescence and the Luminescence of Incandescent Bodies," Carnegie Inst. of Washington, 1928. Publ. No. 384.

[8] Nichols, Howes and Wilber, op. cit., p. 341.

[9] Gibson, K. S., *Phys. Rev.*, [2], viii, 38, 1916.

[10] Tanaka, *J. Opt. Soc. Am.*, viii, 287, 1924.

[11] The beginning of the first step is chosen arbitrarily and determines the beginnings of all the following steps.

THE EFFECTS OF AN HOMOLOGOUS SERIES OF AMINES ON THE MOBILITIES OF IONS IN HYDROGEN GAS

BY LEONARD B. LOEB AND KARL DYK

Communicated December 19, 1928

Introduction.—One of the writers[1] has undertaken a study of the mobilities of ions in gaseous mixtures with the hope of arriving at some clue as to the nature of the gaseous ion. That this study has not been in vain is shown in a recent summary of the results published by one of the writers.[1] It indicates that while in some gaseous mixtures the mobilities follow Blanc's[2] law which indicates no change in the fundamental nature of the ion (its physical radius, not the electrical shortening of its mean free path) in mixing the gases, other mixtures show marked changes. One type of change indicates what may be a sort of change in ionic dimensions, (a statistical clustering); the other the formation of a definite ion due to the rearrangement of the molecule or molecules surrounding the ion caused by an especially great affinity of the ion for one of the type of molecules of the mixture. In such mixtures the first trace of the more active constituent added to a binary mixture abnormally decreases the mobility of the ions, or, using a more convenient form of representation, abnormally increases the resistance to the motion of the ion so that a marked departure of the linear law of resistances in mixtures characterizing Blanc's behavior manifests itself.

One of the most striking results obtained was with NH_3[6] in air and hydrogen where the mobility of the positive ion is increased to nearly the value of the negative ion for quantities of NH_3 of the order of 0.1% or less. This indicates the formation of a less bulky ion (possibly NH_4^+), due to the replacement by NH_3 of some large molecule of impurity to which the positive ion normally attaches itself in air or H_2. In discussing this phenomenon with Professor C. A. Kraus the latter suggested that a study of some of the amines might prove of interest in this connection since in those cases the size of the molecule could be varied without markedly changing the electro-chemical properties of the gas. Measurements were therefore undertaken in mixtures of hydrogen and methyl-amine (CH_3NH_2) and hydrogen and *N* propyl-amine CH_3-CH_2-CH_2NH_2. The chief difficulty in such studies of organic derivatives lies in the difficulty of obtaining

specimens of the amines which do not contain other members of the same series, for in organic chemistry the reactions are such as to give several different products which are so nearly alike chemically as to be exceedingly hard to separate in the pure state where quantities as small as 0.1% of foreign constituents are liable to seriously change the results.

Method.—The methyl-amine was prepared by dropping concentrated aqueous solutions of NaOH onto a commercially pure sample of the hydrochloride CH_3NH_3Cl (Eastman Kodak Co., white label) at reduced pressure, passing the amine over a cooling bath to remove the greater part of the water vapor, than over a meter-long tube of powdered NaOH into a freezing tube immersed in liquid air. After 25 cm.^{3} of the liquid had been collected the generator was shut off and the solid in the tube was allowed to warm up. The first third of the liquid was distilled off and the second third was allowed to pass over another meter-long tube of NaOH into a reservoir at liquid-air temperature in vacuum. The final third was run out through a filter pump. The gas from the reservoir was used in the measurements by warming it up to about 15 cm. or more vapor pressure and distilling it into the ionization chamber as needed. The only substance which came in contact with the amine frozen in the reservoir was mercury vapor from the manometer which gradually over a period of time distilled over into the reservoir. The propyl-amine was taken from a sealed flask of the liquid as sold by Eastman Kodak Co. (white label). It was distilled into a trap using liquid air. From there it was fractionated into a second trap immersed in liquid air, passing over a meter-long tube of NaOH crushed into a fine powder. Thereafter it was fractionated as was the methyl-amine into a reservoir with liquid air.

The measurements were made following the standard practice of one of the writers.[3,4,7] For the measurements in methyl-amine an auxiliary field of 6 volts per cm. was used, while for the propyl-amine an auxiliary field of 5 volts per cm. was used, the gauze, electrodes, chamber[4] and procedure being the same as used in the study of other gaseous mixtures.[3,7]

Results in Pure Methyl-Amine.—In using both amines a great deal of trouble was encountered due to condensation or adsorption of the amine on the metal walls of the chamber. With methyl-amine this produced what may have been contact potential effects on the gold-plated electrodes so that the results at higher pressures were accompanied by disturbances that made measurements above 40 cm. pressure unreliable. Among other effects the curves below the potential where ions cross were characterized by irregular negative readings and the intercepts with the zero axis were such as to make accurate evaluation questionable. The curves at these higher pressures even appeared to indicate a break point suggesting two mobilities. Where the break point occurred in the table of results the values in parentheses give the mobility estimated. As work with mix-

tures later showed it seems unlikely that the ions of different mobilities were really present, the effect being due to irregular deflections near the point where the ions ceased to cross in the period of alteration, possibly due to condensation of amine on the insulating surfaces of the chamber. The mobilities of the ions observed were all corrected to values at 760 mm. pressure by the inverse pressure law.

MOBILITIES IN METHYL-AMINE AS ACTUALLY OBSERVED REDUCED TO 760 MM. PRESSURE

POSITIVE IONS	NEGATIVE IONS	PRESSURE IN MM.
0.482	0.587	686
0.598 (1.99)	0.598 (1.99)	613
0.509	0.576	412
0.516	0.540	256
0.544	0.590	149
Average 0.530	0.574	

On the old accepted standard of mobilities in air as 1.4 cm./sec. for positive and 1.80 cm./sec. for negative ions per volt/cm. the values obtained are 0.494 cm./sec. for positive ions and 0.535 for negative ions per volt/cm., while on the new absolute standard of one of the writers[5] they are 0.590 and 0.640 cm./sec. per volt/cm., respectively. The values in NH_3 as given elsewhere[6] are 0.62 cm./sec. for the positive ions and 0.67 cm./sec. for the negative ion per volt/cm. on the old standard and 0.707 and 0.77 cm./sec. per volt/cm., respectively, on the new absolute standard. The mobilities in the amine therefore are somewhat less than in NH_3, an effect probably due to the larger size of the methyl-amine molecule.

Results in Methyl-Amine Hydrogen Mixtures.—The results in the methyl-amine hydrogen mixture are given in the curves below. In computing these results the average mobilities in the amine *as observed* were used while the mobilities in H_2 used were 6.32 cm./sec. and 9.85 cm./sec. per volt/cm. for positive and negative ions, respectively, uncorrected for the effect of the auxiliary field. The mobilities were largely observed between 20 and 40 volts per cm. alternating field and with an auxiliary field of 6 volts/cm. in order to make them comparable without correction. The curve of figure 1 gives the results for mixtures from 0 to 10% amine, while figure 2 gives the results from 0 to 100% amine. The abscissae give the percentage of amine in hydrogen while the ordinates are in reciprocals of mobility (i.e., equivalent resistance). The two straight lines represent the computed values on the basis of Blanc's law from the values for pure hydrogen and amine, the dashed curve and line being for negative ions, the full curve and line for positive ions. The points and curved lines represent the *observed* values of the *reciprocals* of the mobilities.

The results are very interesting for they indicate that the *negative* ion shows a marked departure from Blanc's law for small percentages of

amine, while the positive ions retain their normal mobility. Experiments with less than 0.25% amine were uncertain due to the adsorption of gas by the walls of the chamber. To remove effectively the amine from the chamber required at least ten hours of pumping and washing with hydrogen several times, by a pumping system which could give an x-ray vacuum in a clean chamber in two hours' pumping. It is also seen that after the initial change in mobility the deviation of the curves for both ions from Blanc's law was small and similar to that for NH_3-hydrogen mixtures.[1,6] The poorer values at higher percentages of amine are due

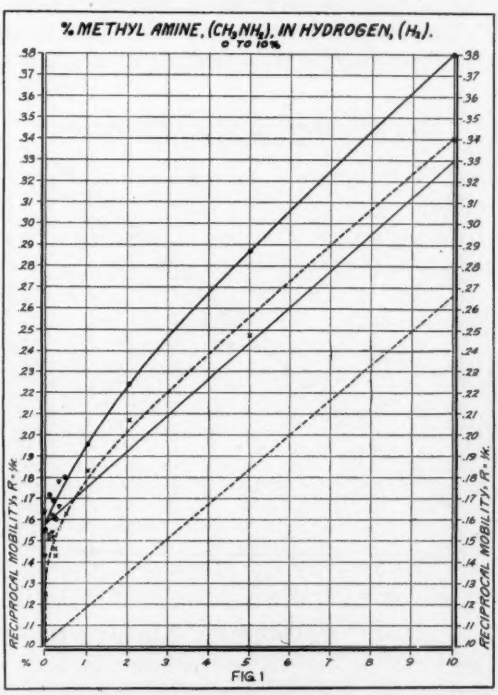


FIG. 1

to the effect mentioneed under a description of the method, that is, the irregular mobility curves for high pressures of amine. That in reality the seeming dual curves indicating two carriers were spurious is shown in these mixture curves by the fact that if mobilities are calculated from the *apparent* break points in the mixture work instead of the extrapolated break points (which would be the case for two classes of ions), the values of the mobilities are very irregular and depart from a regular curve in highly fantastic fashion, which is not characteristic of mobilities in mixtures.

Mobility of Ions in Pure Propyl-Amine.—In the case of the measurements of mobilities in propyl-amine the chance for the presence of higher members of the series and other impurities is much greater than was the case with methyl-amine. Furthermore with this amine the condensation of the amine on metal surfaces could be actually followed. At the temperature of the room (25°C.) the propyl-amine had a vapor pressure of some 280 mm. Measurements of mobility at this pressure were, however, impossible, as was the case even at 130 mm., since the amine had condensed over the polonium source (a copper disc coated with Po and Bi) in such a thick layer that the α-rays could not get through except for the last half-cm. or so of their range of 3.8 cm. in air. At 75 mm., however, the full range of the α-particles emerged through the layer of amine and mobility

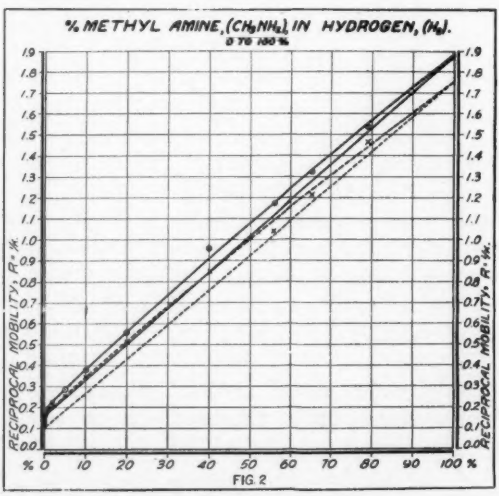


FIG. 2

measurement was possible. The mobilities observed in the amine reduced to 760 mm. were 0.273 cm./sec. for positive ions and 0.288 cm./sec. for negative ions per volt/cm. On the old standard scale for air of 1.4 cm./sec. and 1.8 cm./sec. per volt/cm. for positive and negative ions, respectively, the values in *n*-propyl-amine are 0.243 and 0.257 cm./sec. per volt/cm., respectively, and on the new absolute[5] standard scale the values are 0.294 and 0.312 cm./sec. per volt/cm. for positive and negative ions, respectively. It is seen that the large bulk of the propyl-amine molecules is very effective in reducing the mobilities, for the electrical effects are not far different from those in NH_3 or CN_3NH_2 and probably less.

Mobilities in Propyl-Amine Hydrogen Mixtures.—The results of the mixture experiments are reproduced in figure 3. In the figure as for

methyl-amine the abscissae are the per cent of amine in hydrogen and the ordinates are reciprocals of the mobility. The full curves are for the positive ions while the dashed curves are for the negative ions. Blanc's[2] law is as before indicated by the two straight lines. These curves are calculated from the *uncorrected* values of the mobilities in the pure amine and the *uncorrected* values of the mobilities with hydrogen in the clean chamber which were 6.38 cm./sec. per volt/cm. for positive ions and 8.67 cm./sec. per volt/cm. for negative ions. Once the amine had been admitted to the chamber it was impossible to remove the last traces except by opening the chamber and giving the whole system a vigorous cleaning with cleaning mixture. The amine seemed to stick to the metal and glass walls and distilled all over the apparatus when near its saturation pressure at room temperature.

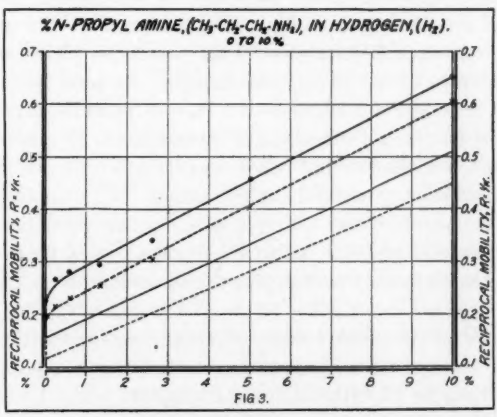


FIG 3.

It is seen that in the case of both positive and negative ions the smallest trace of amine reduced the mobility (increased the reciprocal of the mobility or the resistance) by at least a factor of two. Thereafter the two curves run nearly parallel to the Blanc's law curve, indicating that the values will eventually approach the values in the pure amine as is the case for methyl-amine and hydrogen and for ammonia and hydrogen.

Effect of NH_3 on CH_3NH_2-H_2 Mobilities.—The effect of the addition of ammonia to methyl-amine hydrogen mixtures was also tried. In mixtures of propyl-amine and hydrogen such quantitative experiments were not carried out as the qualitative trials showed that NH_3 did not materially affect the lowering of mobility produced by $C_3H_7NH_2$. The results obtained in CH_3NH_2-H_2 mixtures with varying amounts of NH_3 are shown as triangles on the mobility curves of figure 1. They were results obtained for mixtures of 0.33% CH_3NH_2 and 1.05% NH_3 in H_2, 0.156% CH_3NH_2

and 0.388% NH_3 in H_2 and 0.000173% CH_3NH_2 and 0.00043% NH_3 in H_2, respectively. They are plotted in terms of per cent CH_3NH_2 in hydrogen and it is seen that the points fall slightly above the corresponding curves for the CH_3NH_2 in H_2 alone.

This indicates that the mobilities are decreased slightly more by NH_3 in an amount to be expected from the increased Blanc's law effect of the NH_3 added. The lowering of the mobility is thus sensibly that produced by CH_3NH_2, and NH_3 does not replace the latter in mixtures with hydrogen.

Discussion and Conclusions.—It is seen at once that *N* propyl-amine adds to both positive and negative ions even when present in small traces producing a very great initial decrease of the mobility. This is doubtless due to its bulk. The long carbon chain even in the event of the addition of a single amine molecule to the ion acts to increase the *real* radius of the ion and thus markedly reduces the mobility. The result is just what one would expect for the positive ion. The addition of the amine to the negative ion was a marked surprise to the writers as NH_3-like compounds are not expected to attach to negative charges. In pure NH_3 electrons are probably not attached to molecules of NH_3 at all and in NH_3 the negative mobility in hydrogen is little changed by additions of small amounts of NH_3.[1,6] There is a chance that other organic substances were present as impurities to which negative ions might attach. The analogous behavior of CH_3NH_2 on negative ions where contamination probably occurred in far smaller measure, and the nature of preparation of these compounds, however, makes it appear that the propyl amine really does attach to negative ions.

The behavior of the methyl-amine is still more interesting. Here the addition of a trace of amine does not raise or lower the positive mobility within the limits of experimental measurement, while the negative-ion mobility is lowered by a considerable factor (i.e., the reciprocal of the mobility is increased). This indicates that if the CH_3NH_2 attaches to the positive ion it makes a molecular addition product or a cluster of the same size as the molecule of impurity to which the positive ion in hydrogen normally attaches itself. The negative ion, however, attaches to the CH_3-NH_2 molecule or molecules, and forms a complex of about the same size (the mobilities are nearly equal) as the positive addition product with the amine. Thereafter with increasing amounts of amine mobilities decrease along curves differing comparatively little from the Blanc's law curves. The amount of deviation with higher concentrations of amine is similar to that observed for higher concentrations of NH_3[6] and can be accounted for by the statistical clustering mentioned elsewhere.[1]

We thus see that the three substances NH_3, CH_3NH_2 and CH_3-CH_2-CH_2-NH_2 act as follows on the ions. NH_3 attaches selectively to positive ions replacing some bulkier constituent which usually attaches to positive

ions in hydrogen, making its mobility equal to that of the negative ion in hydrogen or nearly so. The negative ion in ammonia is, however, little affected. In CH_3NH_2 the positive-ion mobility is little changed. It is to be presumed from the behavior of positive ions in NH_3 and $C_3H_7NH_2$ hydrogen mixtures that CH_3NH_2 attaches to the positive ion for NH_3 and $C_3H_7NH_2$ both attach and CH_3NH_2 lies between them in properties. Thus, the fact of no change of positive mobility on addition of CH_3NH_2 to hydrogen can only mean that the addition product of the positive ions to CH_3NH_2 has about the same radius as the normal impurity addition product on positive ions found in hydrogen alone. This conclusion is very important for it delimits the qualities of the unknown addition product on positive ions in H_2 alone. In CH_3NH_2, contrary to its action in NH_3, the negative ion attaches CH_3NH_2 and forms an ion of mobility *nearly as low* as that of the *positive* ion usually observed in hydrogen. Whether the small differences in the two mobilities in the mixtures are due to dielectric effects or to actual differences in cluster size cannot be stated at present. Finally in $C_3H_7NH_2$ both ions form such bulky addition products that their mobilities are both reduced by the largest factor yet observed in mixture studies. The dielectric attractions of these ions on amine molecules again are not quite equal or else the ions may have clusters of different size about them, for the two mobilities in the mixture are not quite the same.

While the results detailed above are quite consistent with the idea that the mobility changes can be described through the formation of addition products involving only one or two molecules with ions (i.e., a small ion theory[1] or the recent theory of Erikson[8]), the writers are not convinced that this explanation is necessarily the correct one. In the first place, the fact that the mobilities of positive and negative ions are never exactly the same, even when the presumption is likely that the addition products may be the same, leaves a doubt, inasmuch as these effects can be explained as conveniently by the assumption of clusters of slightly different size as well as by different residual dielectric attractions.[1] Again some peculiar unpublished anomalous results in SO_2[7] and Cl_2 indicate a possibility of positive-ion mobilities as much as twice as high as any positive or negative mobilities observed in hydrogen before. These results which require careful repetition demanding new features in technique are at present under way. Should such values actually be observed beyond doubt the large cluster nature of the ion would be strongly indicated. Accordingly, while the small ion explanation gives a convenient representation of the results the writers are loath to draw conclusions at this juncture.

Note on the Simultaneous Observation of More Than One Mobility in a Gaseous Mixture under Ordinary Conditions of Measurement.—Had not the peculiar apparent break points in one or two of the methyl-amine mobility measurements been indicated as spurious by the measurements in mix-

tures, the writers would have interpreted these as indicating the presence of two types of ions having different mobilities existing in the gas at the same time. It would have been one of the few cases on record where such mobilities were definitely observed by the method used in a form where the nature of the curves left no doubt of interpretation. The experiments of Haines,[9] Nolan,[10] and Wahlin[11] leave much to be desired in the certainty and significance of their results as some of their breaks are largely discontinuities in what would appear to many like a continuous curve with highly irregular readings. If two different carriers are present curves such as observed by Wellisch[12] and one of the writers[13] indicating free electrons should be observed with this method. With such carriers curves are obtained leaving no doubt as to the presence of two mobilities and the method is sensitive to the presence of as few as ten per cent of one class of carrier in work of this sort.[14] Two of the curves in methyl-amine out of five observed had this dual character to such a marked extent as to leave no doubt. The other curves failed to show the effect and as stated it proved to be spurious. Accordingly, in the long experience of one of the writers with the alternating current method of measuring mobilities in the most diverse gaseous mixtures and gases it has never been possible to observe ions of one sign of more than one mobility in a gas at a given time; although the method would be easily sensitive to ions differing in mobility by 30 per cent if as much as 10 per cent of one of the carriers were present. The conclusion to be drawn is that under the conditions of observation only one positive and one negative ion mobility is observable in a gas at a given time.

Using an air blast method Erikson[15] had observed without question the *transitory* existence of two positive ions in air and some other gases. The change of mobility indicated in Erikson's experiments is affected by water vapor which retards the transition and makes it observable in time intervals of the order of from 10^{-3} to 10^{-2} seconds.[16,17] Thus, *temporarily at least*, ions of two different mobilities of same sign may exist at once. If, however, the ions are simple addition products of charged molecules with certain especially attracted molecules and these, as shown in this paper, by their bulk profoundly influence mobilities one may well ask why two or more ions of different mobilities are not observed in some of the mixtures studied, such as those above in more permanent form.

The answer becomes simple when the time scale of the phenomena involved as well as the concentrations of molecules present are considered. Mobilities of ions are measured under most conditions (except Erikson's) in periods of time of the order of 10^{-2} seconds. In many cases the ions coming from auxiliary fields are even older (i.e., 10^{-1} seconds). Now a gas *molecule* makes some 10^9 impacts with other molecules a second, while an *ion* from its low diffusion coefficient may make even more collisions.

In 10^{-2} seconds ions have collided with 10^7 molecules. If some very active molecule were present in a gas to the extent of 0.01 of a per cent, 10^3 of these molecules on the average would have collided with an ion in the period during which its mobility was being measured. If the union of ions and molecules is at all stable, and the effects of such gases as NH_3 and ether on positive ions show that they must be stable, all the ions would have had an opportunity to acquire their most stable form. If there were a number of molecules present, each forming an ion of stability different from the others, then only the most stable ion would form with the many encounters possible with the molecules in the time intervals used. It is, in fact, exceedingly rare in any technique so far used in the study of gaseous ion mobilities that less than 0.001 percent of some impurity is not present. It then depends on the stability of the different ions and on the numbers of different molecules present in a cm.[3] of gas which value of the mobility will be observed, but it is obvious that one ion of each sign only will be observed even with higher resolving powers for ions of different mobility in a gas.

The only conditions under which two mobilities could be observed for ions of the same sign in a gas at once, as usually studied, would be the following: Assume that there are a considerable number of molecules present in a gas which has a fairly high affinity for one or both ions. In addition, assume that there are present besides this gas a *small number* of molecules of some impurity with a *much higher affinity* for the ions than the first gas mentioned, such that in only one or two impacts of the ion with the latter molecules the more prevalent gas molecules would be shed for the rarer, more active impurity. If just enough of this active impurity were present so that 20 per cent or so of the ions only collided with these molecules before having their mobilities measured, then it might be expected that two mobilities would appear simultaneously. It is readily seen that such an occurrence will be, to say the least, rare with the numbers of molecules of foreign substance in the chemically purest gases and the wide range of affinities between ions and molecules. Thus, in general, the search for the existence of several ions must be made with methods involving very much shorter time intervals than usually worked with. The search after *several classes* of ions in a gas with the present mobility methods is thus futile and it is hardly surprising that all experiments directed to this end have yielded negative or doubtful results.

In conclusion, the writers wish to thank Mr. F. H. Sanders for his kind assistance in the tedious task of making the many mobility measurements required in this work.

[1] Loeb, L. B., *Phys. Rev.*, **32**, 81, 1928.
[2] Blanc, A., *Jour. Physique*, **7**, 825, 1908.
[3] Loeb, L. B., *Proc. Nat. Acad. Sci.*, **14**, 192, 1928.

[4] Loeb and Cravath, *Jour. Opt. Soc. Am.*, **16**, 191, 1928.
[5] Loeb, L. B., *Jour. Franklin Inst.*, **196**, 537 and 771, 1923.
[6] Loeb, L. B., *Proc. Nat. Acad. Sci.*, **12**, 677, 1926.
[7] Du Sault and Loeb, *Ibid.*, **14**, 384, 1928.
[8] Erikson, H. A., *Phys. Rev.*, **30**, 339, 1927.
[9] Haines, W. B., *Phil. Mag.*, **30**, 503, 1915.
[10] Nolan, J. J., *Phys. Rev.*, **24**, 16, 1924.
[11] Wahlin, H. B., *Proc. Nat. Acad. Sci.*, **10**, 475, 1924.
[12] Wellisch, E. M., *Phil. Mag.*, **31**, 186, 1916; **34**, **55**, 1917.
[13] Loeb, L. B., *Phys. Rev.*, **17**, 89, 1921.
[14] Loeb and Du Sault, *Proc. Nat. Acad. Sci.*, **14**, 192, 1928.
[15] Erikson, H. A., *Phys. Rev.*, **25**, 502, 1924.
[16] Mahoney, J. J., *Ibid.*, **33**, 1929 (in press).
[17] Erikson, H. A., *Ibid.*, **32**, 791, 1928.

---

*LATE GEOLOGIC DEFORMATION OF THE APPALACHIAN PIEDMONT AS DETERMINED BY RIVER GRAVELS**

BY MARIUS R. CAMPBELL

DEPARTMENT OF THE INTERIOR, U. S. GEOLOGICAL SURVEY

Read before the Academy, November 19, 1929

Introduction.—One important phase of the work of a geologist is to determine the structure of the earth's crust, including not only the vast movements that have resulted in the upheaval of mountain ranges, but also the minor flexures of the rocks, generally known as anticlines and synclines, that may or may not have appreciably modified the surface features. The recognition and determination of these minor features in a region composed of horizontal rocks is a comparatively easy matter, but in a region such as the Appalachian Piedmont, where the rocks are extremely old and have suffered untold periods of deformation and metamorphism, the geologic structure is generally so complicated that it completely obscures any minor folds or inclined surfaces that may have been produced by recent warping.

Because of this condition, geologists have not been able, so far, to determine whether or not the Piedmont belt, extending from central Georgia to the Hudson River, has been deformed by recent crustal movements, although the unusual altitudes of 900 to 1100 feet in this belt in Maryland and southeastern Pennsylvania has suggested such an origin.

Some years ago the writer conceived the possibility of discovering slight flexures in this region by careful tracing of high-level gravels which are known to occur near the mouths of the rivers crossing this belt. He found on the Potomac some indications, but no direct proof. No opportunity

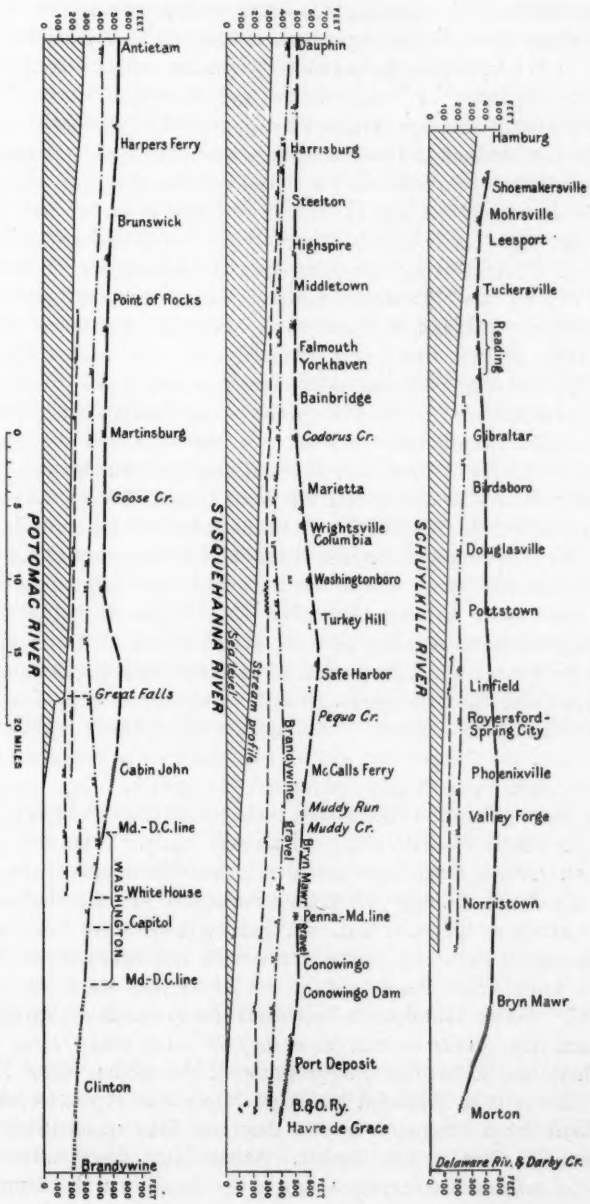


FIGURE 1

Profile sections along three rivers of the northern Appalachian Piedmont region, showing deformation of gravel beds.

presented itself of extending this work to other rivers until the past summer when, through the coöperation of Dr. George H. Ashley, State Geologist of Pennsylvania, he was able to examine with considerable care the gravels deposited by Susquehanna and Schuylkill rivers. The former afforded fairly complete evidence in support of his hypothesis. In this work study was confined to two gravel beds, shown in figure 1, which have been recognized and mapped near the mouths of the rivers mentioned above.

The exact geologic ages of these gravel beds have not been determined, but the higher and older bed is generally recognized as having been deposited in late Tertiary time, but long antedating the Pleistocene ("Glacial") epoch in which similar material was brought into this region from the north. This bed is called Bryn Mawr because of its occurrence in the town of that name near Philadelphia. The age of the lower and younger bed of gravel is somewhat uncertain, but it was probably deposited in the earliest part of the Pleistocene or the closing stage of the Tertiary; it is named Brandywine, because of its exposure near the village of that name in Maryland about 15 miles southeast of Washington.

Susquehanna Section.—Near Havre de Grace, Md., at the mouth of the Susquehanna, the Bryn Mawr gravel caps the upland on both sides of the river for a distance of 5 miles, rising inland from an altitude of 400 feet near Havre de Grace to 480 feet at the upper limit of its exposure. Beyond this point for a distance of 9 miles there are no outcrops because, since its deposition, erosion has been so extensive that all traces of the gravel have been removed. On account of this condition, it was found necessary to trace also the Brandywine gravel, which shows near Havre de Grace at an altitude of 300 feet. Although the Bryn Mawr gravel rises steeply upstream, the Brandywine gravel was found to be nearly horizontal, but on the whole with a slight rise toward the interior.

Figure 1 shows the determined positions of the two beds of gravel and also the points at which observations were made. From this figure it will be seen that the Bryn Mawr gravel, which at the mouth of the river is 100 feet above the Brandywine gravel, rises until at Safe Harbor, Pa., it is 270 feet above the same bed, and that northwestward from Safe Harbor, it changes its slope and gradually descends upstream until at Middletown, a few miles below Harrisburg, it is only 70 feet above the Brandywine gravel. Above Middletown both beds rise gradually upstream and show no such discordance in interval as they do below that place.

There can be but one interpretation of the section below Middletown, and that is that, although the Bryn Mawr was deposited with a gentle gradient down stream, it is now deformed into an anticline whose axis crosses the river at Safe Harbor. As the later gravel shows little or no trace of deformation corresponding to the anticline in the Bryn Mawr, it is positive evidence that the deformation occurred before the younger gravel

was deposited. This does not fix the date of folding very definitely, but it seems probable that additional study would result in determining the geological date of this slight uplift within rather narrow limits.

Potomac Section.—The Susquehanna section makes it possible to interpret the Potomac section with considerable confidence, as the similarity is quite remarkable. In the Potomac section, as shown in figure 1, the Bryn Mawr gravel is well exposed near the Fall Line, and it rises toward the interior from an altitude of 400 feet on the Fall Line to 520 feet in Virginia just south of Great Falls. Beyond this point the land is too low to retain traces of the gravel and consequently for a distance of 11 miles there are no exposures by which to determine its continuity and position. The next exposure is at Daysville, Va., where the gravel caps a hill at an altitude of 430 feet.

In previous work the writer was not able to determine to his satisfaction that the gravel here at 430 feet is the same as the gravel at 520 feet, and if it were the same, whether they were connected by a curved surface or were faulted and failed to connect. The finding, however, of an unbroken anticline on the Susquehanna in approximately the same position relative to the Fall Line makes it reasonably certain that on the Potomac the Bryn Mawr gravel is bowed upward into an anticline whose axis crosses the river just above Great Falls and whose height on the axis is about 100 feet above the same gravel bed on either flank of the fold.

In order to be certain of this interpretation, the Brandywine gravel, as shown in figure 1, was traced from its type locality at the extreme right end of the section, where it has an altitude of 200 feet, up to 350 feet in the western part of the city of Washington, and thence almost continuously up the river gorge to Daysville, Va. Here it has an altitude of 320 feet with the upper gravel 100 feet above it. As in all cases the Brandywine gravel bed is the first one below the Bryn Mawr, it seems reasonably certain that the upper gravel at Daysville is the Bryn Mawr and that if intermediate exposures of this gravel could be found the anticline would be complete and similar in all essential respects to the anticline on Susquehanna River.

In the Potomac section the Brandywine gravel also shows some irregularity which may be interpreted as a slight fold, but as there is some doubt about the meaning of this irregularity, it will not be considered at the present time.

Schuylkill Section.—Much of the land along the Schuylkill River is too low to retain even fragmental deposits of the Bryn Mawr gravel, and hence this section adds little to the solution of the problem, but it is subject to an interpretation which is at least not out of harmony with that given to the other sections.

Between Norristown and the mouth of the river in Philadelphia the

Bryn Mawr gravel is exposed for a distance of about 7 miles and it rises from an altitude of 200 feet at the lowest point to 460 feet at the town of Bryn Mawr. Beyond Bryn Mawr the land adjacent to the river is too low to have retained any trace of this gravel bed, but in the city of Reading it was found at an altitude of 360 feet. This identification is not positive, as the exposure is poor, and the exact altitude could not be definitely determined from the map, owing to recent building operations.

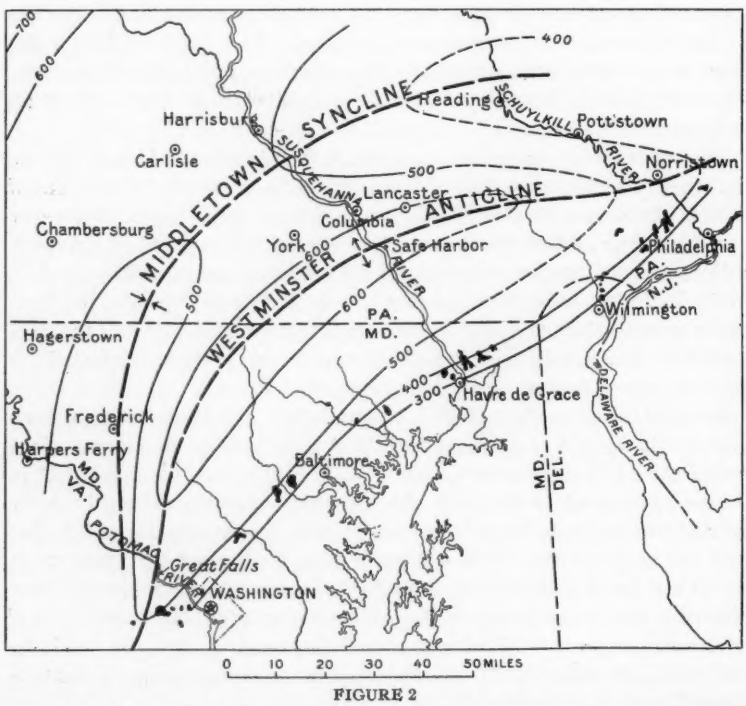


FIGURE 2

Map of eastern Maryland and Pennsylvania, showing by contours the deformed surface of the Bryn Mawr (Tertiary) gravel. Contour interval, 100 feet; Datum, mean sea level. Large areas of Bryn Mawr gravel shown in black.

Above Reading this gravel was followed without difficulty. It was found in good condition about 5 miles above the city at an altitude of 340 feet, which appears to be the lowest point of the gravel on the upper part of the river. Above this point it was found at Mohrsville at an altitude of 370 feet and above Shoemakersville at about 400 feet. The identification of this bed rests on the tracing of the Brandywine gravel from Norristown up the river to Reading.

The Brandywine gravel was not certainly recognized south of the town of Bryn Mawr, but it is well developed about Norristown at an altitude of 200 feet. From this place it was traced almost continuously up the river to Reading, where it attains an altitude of 250 feet. This tracing was verified by the tracing up the river as far as Pottstown of a well-marked bed of gravel about 60 feet below it. The last mentioned gravel is undoubtedly of Pleistocene ("Glacial") age and younger than the Brandywine gravel, which lies above it. Because of these facts the writer feels confident that the Brandywine gravel is present at Reading at an altitude of 250 feet and that the gravel higher in the hills is the Bryn Mawr. If this interpretation is correct, Schuylkill River is crossed by an anticline probably near Norristown and by a shallow syncline a short distance above Reading.

Conclusions.—As each of the three river sections shown in figure 1 affords either positive evidence or suggestions of crustal deformation, between the time of deposition of the Bryn Mawr and the Brandywine gravels, it now remains to plat this fold on a map of the region to see whether or not it agrees with the known facts between these streams. This map is shown as figure 2. The axis of the anticline, which for convenience is called the Westminster anticline, and that of the syncline, called the Middletown syncline, are indicated by appropriate symbols and their positions where they cross the Susquehanna and Potomac Rivers may be regarded as fixed, but the points where they cross the Schuylkill are uncertain, owing to the lack of definite evidence. The form of the warped surface of the Bryn Mawr gravel is indicated by contour lines 100 feet apart. As thus drawn, all of the lines correspond very closely in direction with the general curve of the Appalachian structures to the northwest, and the axis of the Westminster anticline coincides with the highest land in eastern Maryland and approximately with that in this part of Pennsylvania.

* Published by permission of the Director of the United States Geological Survey.